NBII-SAIN FY05 ROAN MOUNTAIN PROJECT FINAL REPORT

Final Report 8/14/06 Version

Thomas Burley, Nora Schubert, Judy Murray, John Peine and Mary Thompson

Table of Contents

Summary

The central objective of this project is a direct reflection of John Mosesso's introductory comments at the March 13-14, 2006 SAIN Partners Meeting: "At the outset, the reason for bringing about NBII was that there were 'gobs' of data and information scattered all over the U.S., not accessible, in incompatible formats, and that NBII was tasked with addressing this problem….NBII's focus is to pull data together that truly matters to someone or communities. Essentially, the core questions are: 1) what are the issues, 2) where is the data, and 3) how can we make it usable and accessible?"[1] This large-scale problem and the related biological issues such as unsustainable resource use and climate change, in addition to effective adaptive natural resource management are also addressed in the Strategic Plan for the U.S. Geological Survey Biological Informatics Program: 2005-2009. "Though the Federal government invests more than $600 million per year in biological data collection, it is difficult to address these issues because of limited accessibility and lack of standards for data and information…..Variable quality, sources, methods, and formats (e.g. observations in the field, museum specimens, and satellite images) present additional challenges." This is further complicated by the fast-moving target of emerging and changing technologies such as GPS and GIS, as well as computing capabilities. "Even though these technologies offer new solutions, they also create new informatics challenges."[2]

The Roan Mountain Project is an excellent example of the value added from the application of NBII expertise and capabilities. The concept of the Data Management Toolkit, (hereafter referred to as the 'Toolkit'), devised for the Roan Mountain Massif (hereafter referred to as 'Roan') can serve as a model for partners and other entities in the context of exactly what John Mosesso describes above: Helping to make data and info that matters available, in a usable format, to people that care. It helps by addressing the broader issue of data integration, which is reliant upon good data management. The Data Management Toolkit can help facilitate better decision making and achievement of goals and objectives by helping to define feasible objectives along with the critical aspects of data that will ultimately contribute and lead to the achievement of those goals and objectives.

Components of the Toolkit include data format, principles of data management, metadata documentation, geo-referencing and data acquisition guidelines, quality assurance/quality control, recommendations for data storage, access, archiving, and custodianship. The devised Toolkit was tested by applying the metadata component to two priority legacy databases collected on the massif related to adaptive management of the grassy balds: 1) transects and plots recording plant community information and 2) mapping geo-referenced boundaries of adaptive management activities (mechanical mowing and hand cuttings). These datasets were ideal test cases in that 1) the transects and plots database incorporated numerous datasets collected over several years by a variety of investigators and 2) the maps were created over a period of 15 years, showing the spatial extent of mowing activities. The latest 4 years also incorporate GPS technology so as to demonstrate the applicability of this crucial tool. Using the process of metadata creation as an evaluation of the datasets, it was determined that the potential to dramatically increase the value of these two example databases existed had the Toolkit been applied from the beginning. This is not to suggest that the efforts were necessarily flawed but rather that the data would have been more useful if 1) more thoroughly and systematically documented and 2) if more specific objectives and subsequent data elements had been defined in the context of those objectives. In addition, application of advanced technology as applied to future data collection can dramatically improve documentation, geo-referencing, archiving, quality assurance/quality control, and appropriate data policy.

[1] Mosesso, John. Address. NBII SAIN Strategic Planning Work Session. University of Tennessee Conference Center Building. 13 March 2006.
[2] Ruggiero, Michael, Marcia McNiff, Annette Olson, and Ben Wheeler. 2005. Strategic Plan for the U.S. Geological Survey Biological Informatics Program: 2005-2009. U.S. Geological Survey, Biological Resources Discipline, 20 pp.

Justification

This project addresses the national goal of NBII and SAIN to develop tools that serve critical data and information management needs. The project addresses a long-standing critical need to provide a platform for the integration of geo-referenced data collected in conjunction with historical and ongoing land/resource management activities across Roan. Although the locations of various species and communities of conservation concern that are present on Roan have been documented by state Natural Heritage Programs, the locations, timing and nature of various land/resource management activities that have been conducted across Roan have not. The inability to effectively collate documentation for and interpret the effects of management activities is in part a reflection of the diverse consortium of land owners and managers who work within this ecosystem. Landownership includes holdings by the USDA Forest Service, the state of Tennessee, the Southern Appalachian Highlands Conservancy and The Nature Conservancy. Management of this landscape involves each of these partners as well as the Appalachian Trail Conservancy, the U.S. Fish and Wildlife Service, the North Carolina Wildlife Resources Commission, the Tennessee Wildlife Resources Agency, and the North Carolina and Tennessee Natural Heritage Programs. This project serves as a much needed prototype methodology applicable throughout the highlands of the Appalachian Mountain range.

Roan is a hotspot of endemic, rare, threatened and endangered species in the Southern Appalachian Highlands, a region of convergence of northern and southern species. The rarity of this ecosystem is reflected in the species it contains: seven are currently listed under the U.S. Endangered Species Act, and an additional species (the peregrine falcon) was declared recovered and de-listed in 1998, though it remains on the Watch List. Another 32 species found on Roan are considered Federal Species of Concern by the U.S. Fish and Wildlife Service (Carolyn Wells, personal communication-2005) and 31 of the bird species found on Roan are in the highest priority categories of the National Partners in Flight Bird Conservation Plans. Roan also contains two of the most endangered communities in the continental U.S (grassy balds and red spruce-Frasier fir forests). As a result of these factors, the U.S. Fish and Wildlife Service considers Roan a regional conservation priority.

The area is also a priority for the USDA Forest Service. In a 3/3/05 email from Terry Seyden, USFS Region 8, National Forests in NC, Supervisor's Office, he addresses the significance of the area and the need for this project:

> "Roan Mountain is of national significance for its natural communities, geomorphic features, and plant and animal species. It is one of the most botanically diverse, higher elevation mountain complexes in the Southern Appalachians. A number of rare plants occur at the Roan Massif. Vegetative diversity along with climatic conditions, topography and geologic conditions provide for an exceptional diversity of animal species, including mammals, birds and amphibians many at or near the southern extent of their ranges. In general, the ecological and geological attributes of the Roan Mountain Massif are of great scientific interest; the aesthetic and recreational popularity of this natural area are unsurpassed and the research and environmental education potential offered by the Roan is unexcelled in the Southern Appalachians.
>
> Improving documentation of and access to biological data collected on the Roan has long been a high priority of forest biologists and managers. We welcome the involvement of the NBII SAIN expertise in addressing this long standing need."

It is widely recognized that restoration and maintenance of the high elevation grassy balds of Roan will require long term vegetation management and monitoring to control the invasion of woody plants so as to preserve its unique ecosystem. This recognition led to the signing of an Environmental Assessment (EA) by the U.S. Forest Service, which prescribed various management activities for the restoration and management of these communities. Implementation of the management techniques described in this EA was preceded by the collection of baseline data on plant community composition along a series of transects and plots established (in the late 1980's) across the grassy balds of Roan. Historically, management and monitoring of the Roan grassy balds has occurred in a piecemeal manner in cooperation with the Forest Service, as funding would allow. There has not been a way for the managers to assess the effectiveness of treatments applied to the area based on documented biological data analysis, nor through monitoring.

In addition to these baseline data collected in the late 1980's, there are other inventory and monitoring data that encompass (in whole or in part) the grassy balds and may inform the management of these resources. These data have been collected over a period of several years by various individuals, organizations and agency partners on Roan. With the exception of locality data on species tracked by the North Carolina and Tennessee State Natural Heritage Programs, these data vary widely in their current format (e.g., electronic or hardcopy) and levels of accessibility to current partners that regularly conduct or advise natural resource management activities on Roan. These partners include the USFWS, the USDA Forest Service, the Southern Appalachian Highlands Conservancy, The Nature Conservancy, the Appalachian Trail Conservancy, and the NC and TN Natural Heritage programs. Since the Appalachian Trail runs through the grassy balds, the National Park Service Appalachian National Scenic Trail Park Office also has a vested interest in the oversight of Roan Mountain. The need for a common repository of management data has been recognized for over a decade (Carolyn Wells, personal communication-2005; documentation and correspondence on this issue available in USFWS files). However, realization of this goal has failed largely because there appeared to be no mechanism for the collation and exchange of spatial data in a dynamic, electronic format. The timing of this project is appropriate to implement a long-term vision. Judy Murray, Highlands of Roan Stewardship Director of the Southern Appalachian Highlands Conservancy sent the following email on September 12, 2005:

> "On March 29, 1999, Nora Murdock, Rob Sutter and I coordinated a meeting of all the partners who share management responsibilities and/or interest in the Roan Mountain Massif. While it may seem that the movement for ecosystem-level and adaptive management planning has ebbed and flowed over the years, pieces of it continued to manifest in a number of ways--including refinement of the Roan Management Area planning in the Cherokee Forest Plan Revision, funding exploration meetings (with the Roan Partners Steering Committee, Cherokee & Pisgah Forest Supervisors and District Rangers), expanded balds management initiatives, GPS tracking of treatment areas and field assessments.
>
> This Data Management Toolkit project on which we are embarking has been a long time in coming, but with advances in technology, the timing is absolutely right. I see it as the key to the future of collaborative ecosystem management for the massif and a model to serve in a global capacity. I thank each of you for the commitment and possibility that you bring to the venture."

The greatest challenge for this project is to design a data management methodology that is flexible enough to handle the variety of data being collected and be utilized by all of the partners involved. Gaining insight on how best to make the information adaptable and accessible to all partners may be the greatest benefit of the project to NBII.

Partners

Organization	Person	Responsibilities/Area of Interest
USGS-SAIN		
	John Peine	Project management
	Tom Burley	Design, implementation and testing of Toolkit
USFWS		
	Carolyn Wells	Threatened and endangered species stewardship, botanist, metadata review
USFS		
	Steve Simon	Forest Ecologist/Botanist, Threatened and Endangered Species Program Manager, National Forests of North Carolina
	Paul Bradley	District Ranger, Appalachian District, Pisgah National Forest
	David Danley	Botanist, Appalachian District, Pisgah National Forest
	Terry Bowerman	District Ranger, Nolichucky/Unaka District, Cherokee National Forest
	Joe McGuiness	North Zone Wildlife Biologist, Watauga District, Cherokee National Forest
SAHC		

	Nora Schubert	Document metadata for legacy databases, collect key documents, conduct literature searches
	Judy Murray	Highlands of Roan Stewardship Director, metadata review
ETSU		
	Norma Myers	Director, ETSU Archives of Appalachia
ATC		
	Matthew Davis	
	Julie Judkins	Program Director
	Morgan Sommerville	Regional Director
NCNHP		
	Misty Franklin	Botanist
	Mike Schafale	Community Ecologist
NCWRC		
	Chris Kelly	Nongame Program Coordinator
NPS		
	Kent Schwarzkopf	Natural Resource Specialist, Appalachian National Scenic Trail
TNC		
	Beth Bockoven	Southern Blue Ridge Coordinator, Mountains District, Conservation Program
	Este Stifel	Director of North Carolina Hickory Nut Gorge Landscape Project, Mountains District, Conservation Program
TNNHP		
	Lisa Huff	Stewardship Ecologist-Upper East Tennessee
	David Withers	Zoologist
	David Lincicome	Rare Species Protection Program Administrator
	Claude Bailey	Botanist
	Andrea Bishop	Rare Species Biologist
WFU		
	Peter Weigl	Professor of Biology, balds origins, small mammals, Northern Flying Squirrel, conservation biology
Biological Survey Contractor		
	Jamey Donaldson	Botanist, threatened and endangered plants

Objectives

The overarching goal is to create a data management methodology for inventory and monitoring activity in support of adaptive management on Roan which will:

1. Define and prioritize data and information elements that affect the ability to monitor and manage the biological resources of Roan.
2. Create for the subject geographic area a Data Management Toolkit based on existing content by facilitating the use of existing infrastructure by land managers and interested parties of Roan. The Toolkit will thereby be utilized to assist in interpretation and analysis of land management activities in this sensitive ecosystem.
3. Identify appropriate databases and evaluate their potential to have the Toolkit applied to them.
4. Test the devised methodology by entering a variety of types of information from identified databases in the context of resource management actions aimed at conserving the high elevation species and biological communities of Roan.

The project began with a meeting entitled “Roan Mountain Massif Information System-First Step in new NBII SAIN project” which was held November 17, 2004 during the annual SAMAB Conference. Representatives of USFWS, NPS, USFS, ATC, TNC and SAHC were in attendance. All agreed that data management related to Roan was a long overdue priority need. The consensus was to focus on the grassy balds and include biological data related to inventory and monitoring and adaptive management of the balds.

Desired Capabilities of Toolkit

The list of desired capabilities requested includes the following:

1. Standardized Format and Guidelines for Data Recording and Georeferencing While in the Field Using USFS, Industry and Federal Standard Procedures and Guidelines

Much of the work done on Roan up to this point has been in various formats, levels of accuracy, and in general fairly inconsistent and incompatible with one another. A clear need for specifications on fundamental elements associated with data collection in the field was identified because of this.

Following standard formats and guidelines will greatly enable the compatibility and utility of data collected on Roan in the context of the common goal of managing and maintaining the balds and the unique biodiverse ecosystem found there. This can help promote consistency and interoperability among datasets pertaining to Roan via a base standard of elements. Also, this helps ensure the integrity of data and information resulting from a project, in essence providing an aspect of Quality Assurance and Quality Control. Levels of accuracy will be more consistent within and across projects, ultimately allowing for better analysis of present conditions, better information, and ultimately better decision making in general as applied to active management. Furthermore, data that is acquired via standard methodology and guidelines allows for measurement of impacts resulting from implemented on-the-ground decisions because the effectiveness of planning and decision-making are closely related to the quality, consistency, and completeness of available information.

2. Metadata Documentation Following Federal Guidelines

Proper documentation is part of the foundation of good data management. Any sort of standard metadata documentation for data resulting from work done on Roan has been virtually nonexistent. Following a common, accepted format for metadata documentation will allow for consistent, dependable documentation for projects and datasets. Frequently, metadata is viewed as an after-thought of projects, a resource that is created once a project is done, if created at all. In reality, proper documentation is central to viewing data and information as a long-term investment in that all too often, real world issues such as employee turnover, equipment change and/or failure, advances in technology, and simply time can result in important details about data being lost, which in turn may determine whether or not a dataset can even be used or compared quantitatively 5, 10, or however many years down the road.

A thorough, standardized documentation methodology helps promote efficient use of resources (time, money) as well as helping to avoid data duplication. In addition, data can be made ‘discoverable’ by other researchers by having data documentation in a consistent, standard format available in existing clearinghouses that can be queried by anyone that uses spatial data.

3. Quality Assurance and Quality Control (QA/QC) Procedures

Quality Assurance (QA) of data as an all-encompassing management approach internally within a project can help ensure that the objectives of the project can be fully met, and that the data will meet the user’s expectations. Quality Control (QC) serves as a mechanism to evaluate and ensure the integrity of a final product and that it meets the agreed upon parameters/acceptance criteria specified in the QA plan. Having a common well-defined understanding of QA and QC and how it may be applied to spatial data can help ensure the overall quality of work and the final product.

In moving towards integrated efforts for adaptive management on Roan, this is a key aspect so as to ensure good data and good information that will ultimately drive adaptive management decisions via defined objectives. QA and QC can also help ensure that people do not have misconceptions about the data, and if they are unsure what their expectations should be, it can help define those for them so that they better understand the specifics of the final product. QA/QC contributes to utilizing resources to their fullest, resulting in a quality deliverable that meets the user’s expectations and that can be relied upon and utilized in the long-term.

4. Archiving Procedures and Specifications for Paper and Digital Records

A major issue with work done on Roan over the years is that much of it has been housed at an individual's home and/or at their respective individual's or agency's office. No real archiving or organizational methodology has been used besides each individual's own internal knowledge of what exists within their stacks and boxes of papers, and as a result there is limited knowledge among partners regarding what fellow Roan stewards may have in their possession. By establishing a centralized methodology and agreement on how relevant materials may be archived, the data and information relevant to Roan can be preserved for future work beyond the individuals whom currently contribute their efforts. In turn, this helps better support the idea of data and information being a long-term investment that requires proactive management and upkeep.

5. Web-Enabled and When Appropriate Controlled Access to Digitized Data and Reports

Web access to spatial data pertaining to Roan has essentially been nonexistent up to this point. By maintaining a centralized Open GIS Consortium (OGC - http://www.opengeospatial.org/) compliant digital repository for geospatial data, a common database pertaining to adaptive management work on Roan can be utilized by the various partners and agencies that contribute to Roan's biodiversity management needs. This can help resolve the issues of disparity among datasets that has been a problem up to this point, and at the same time can allow all partners access to relevant data from other partners and agencies so as to take advantage of existing data that may possibly meet a need that they have.

Much of the data contained in the Roan legacy databases has at least general references to Threatened and Endangered (T&E) species due to the unique nature of the ecosystem there. As a result, controlled access to such locational data, even if simply for general portrayal purposes, has been a concern voiced by the partners involved. Where appropriate, subsets of data can be created and made available that are determined acceptable in regards to sensitive content. In addition, additional security parameters, depending on how the data and information is made accessible, can be implemented such as password logins, etc.

6. Permanent Collection Facility for Storage of Data, Reports, Articles and Management Records

During a meeting on January 24, 2006 at the Sherrod Library at East Tennessee State University in Johnson City, Tennessee, the decision was made to establish ETSU's Archives of Appalachia as a central repository for natural and cultural information pertaining to Roan Mountain. A permanent collections facility for Roan-related management and policy reports, scientific papers, and original data has been identified as a need due to the current scattered state of materials pertaining to Roan's adaptive management history, a need which has been recognized by federal, state, private and academic partners of the Southern Appalachian Highlands Conservancy Roan Stewardship Committee over the past several years. This centralized repository will help directly address the issue of access for the various partners, as well as the important aspects of proper archiving and custodianship of data and information. By identifying and engaging an institution with the capacity to properly manage such materials such as ETSU's Sherrod Library, this institutional commitment will help anchor the probability of sustaining commitment to the information management system being developed in the long term.

7. Future Focus on How Information Collected Will Be Relevant to Scientific Inquiry

By addressing the latter identified needs in the context of data and information pertaining to Roan and its unique grassy balds, legacy data and information will be better suited for scientific analysis and inquiry as applied to adaptive management. Furthermore, by following such protocols for work done on this specific landscape in the future, the stewardship community of Roan will be better able to utilize new data and the resulting higher quality information so as to make better informed decisions regarding Roan's adaptive management.

Toolkit Components

The components of the Toolkit consist of the following:

1. Fundamentals of Data Management: Basic Elements, As Well As Relevant Questions and Considerations in the Context of Data Management

The basic elements of good data management and how to approach such an important part of any project are discussed. A brief overview of the importance of standards, an overview of the elements that make up data management, strategic implementation tips for projects, important considerations in the context of data acquisition, as well as the idea of thematic data content standards are given. This component of the Toolkit will serve as a higher-level guidance tool for incorporating proactive data management concepts into projects and agencies focused on a common goal.

These important aspects of data management all contribute and increase the probability of long term relevancy of the data in the context of adaptive management and science. In addition, it will enhance efficient use of project and organization resources (time, money) in both the short term and long term, and can ultimately help address the larger issue of data integration across projects and partners involved in a common effort. Standards such as thematic data content standards can be used as a base minimum of attributes that should be included with data so that types of data can be integrated with other types that also have the same minimum attributes. The rest of the Toolkit conceptually crosswalks back to the elements of data management and the strategic concepts presented in this section.

2. Spatial and Non-Spatial Data Elements

Being cognizant of the aspects of spatial data, as well as elements common to data in general all should be considered when a project is being planned. Fundamentals spatial elements such as scale, datum and projection, as well as non-spatial elements such as file naming conventions are addressed and described so that Toolkit users can better understand these elements in the context of their projects and data. All of these ultimately influence the integrity of data and information as well as the ability of data to produce information needed to meet a project's objectives and goals. Geo-referencing procedures in the field are evolving rapidly due to the elevation of GPS technology and GIS related software. It is extremely important to document the hand-held instrumentation in use and the software applied in analysis. Resolution and accuracy is a moving target.

3. Quality Assurance and Quality Control (QA/QC) For Spatial Data

The idea of Quality Assurance (QA) is presented and demonstrated as to how a QA Plan component of a project truly provides an all-encompassing management approach internally for a project. By integrating every aspect of a project so as to fully meet the expectations set for the end product, a QA plan can help streamline work and use of resources. The QA component of the Toolkit covers the basic concepts of QA and provides suggestions as to how such an important management tool can be integrated into projects resulting in spatial data.

Quality Control (QC) is also discussed, and some suggested methods for applying QC to spatial data are presented utilizing ESRI ArcGIS software. By establishing an agreed upon QA plan with specified QC techniques and confidence levels, misconceptions about data can be avoided. Also, if the end users are not sure of what their expectations should be, it can help better define those for their understanding.

4. Data Structuring and Modeling

Development of a data model is a crucial step of the planning process which can help ensure that all aspects of the real-world system are represented, as well as helping to ensure that the goals and objectives will be met by cross-walking the elements of the real world system to elements of a database. Integrating a model into a GIS environment can be beneficial in that a geodatabase model will help bring a physical data model closer to the logical data model that it is based on. By establishing this architecture framework at the outset, problems that could possibly derail a project (can be avoided down the road. Some risks of not properly addressing this step include poorly organized and structured data, incomplete data that does not meet the needs and goals of the project, duplicate, missing, or unnecessary data, bad representation of data, and lack of proper data management implementation relative to the data associated with the project.

For projects with data that will require spatial analysis, querying, and cartographic output, data modeling should be explored so as to better support these functions. The concepts of data modeling and organization are presented with a three-tiered approach in terms of a Conceptual, Logical, and Physical model workflow so as to create a database structure that functions as a representation of a real-world system in a GIS format. Utilizing this approach for the structuring of data will ultimately allow for better cartographic output and overall data analysis.

5. Metadata Documentation Following the FGDC-NBII Standard Format

A series of questions designed to extract relevant information needed to generate an FGDC-NBII (Federal Geographic Data Committee - National Biological Information Infrastructure) compliant metadata record is being used. A crosswalk is included which shows how the questions relate to the FGDC standard. Similar questions are included for utilizing the NBII endorsed Dublin Core Standard for documenting web-enabled resources as well as reports and documents. Included is a step-by-step 'suggested workflow' outline for Toolkit users as well as a list of free software currently available for facilitating creation of a compliant record.

No metadata records existed for any of the legacy databases being documented during FY05, so as an applied demonstration metadata creation has been done for two of the identified priority legacy databases (further discussed in this report). The project partners are lacking in metadata capabilities, expertise, and standard dataset documentation protocols, a critical issue with which this methodology will help assist. And most importantly, there is very little appreciation of the importance of this function. This project dramatically illustrated that importance as to the value to the evaluated legacy databases.

6. Suggested Guidelines and Procedures for Spatial Data Recording and Georeferencing Both in the Field and in the Lab

Utilizing standardized formats and specifications for spatial data collection will allow for consistency in data accuracy as well as compatibility among data sets for interoperability and integration. Georeferencing of field work via GPS technology following USFS, industry, and Federal guidelines and standards are included in a data acquisition section in as straightforward a manner as possible for practical use by those working with spatial data. This topic is approached from the perspective of common issues and decisions that must be made in regards to equipment types, data collection parameters, and issues associated with spatial data collection that individuals might not be aware of or may not fully understand. This Toolkit component is designed to cover elements that are not specific to one particular brand or model of natural resource grade GPS unit, but does assume access to at least ESRI ArcGIS software for processing and analysis.

In addition, common issues that arise with data acquisition from other sources are addressed with such things as scanning and georeferencing, as well as heads-up digitizing. A brief overview, items of consideration, and if applicable suggested process steps are provided for the latter data acquisition methods. Working with 'tabular' data is also briefly overviewed with basic principles and concepts related to integration into a GIS.

7. Data Storage, Access, and Custodianship

In keeping with the approach of viewing data as a long term investment, an appropriate data policy should be created that can help manage the data once a dataset is completed for post-production lifecycle control. The data will need to be made accessible to people and researchers, and depending on the type of data, may require restricted access. The data will need a place to live (storage) and the appropriate 'climate' that correlates with the data's importance and needs. The data will need specified procedures in place should updates need to be made. The data will need a 'caretaker' or custodian who can be made accountable and whom can handle the responsibility for the management of the data over the long haul. This component of the Toolkit addresses these important aspects of a dataset's lifecycle by giving an overview description of how each element relates to a dataset, and where possible provides considerations and examples of how each element may be addressed. At the very least, it serves to make users aware of these critical elements related to data.

Having the latter elements clearly defined, possibly as a series of two or three various standard agreed upon scenarios that can be readily applied to any future work, can greatly increase the utility of data in the long term. This can then ultimately result in economies of scale by spreading out the cost of a dataset by making it fully usable beyond the initial need for the data.

A good model for access to sensitive data such as T&E species are those utilized by the Texas Parks and Wildlife Department[3].

> "Permits are required to access the data but access is encouraged. Related to this issue, all data were digitized at TPWD on their computers thereby, all involved in the process will be working from the same data sets.
>
> The issue of masking point location specificity is a bit complicated. NatureServe uses polygons for all types of data but like Texas, when appropriate, provides fairly precise site locations. The buffering strategy all depends on the application of the data. A key goal is to make the information as useful as possible to the user. Only three people (all TPWD employees) are authorized to directly access and manipulate these data as needed. All other access is restricted by fire-walls. The point data are masked by a polygon so that the actual latitude-longitude coordinates are not divulged. The polygons are typically 1 mile wide."

8. Guidelines for Cartographic Output

Utilizing appropriate techniques for display of information resulting from data and analysis can be, if done correctly, an efficient means of conveying a large amount of information in a condensed format to a specific audience. Elements such as the basic parts that make up a 'good' map, as well as guidelines and suggestions for display of various types of information are included in this component.

Specifying these formats and output details for the types of information that is anticipated to be extracted from data in the form of maps can help the end users of datasets better understand the resulting information. In addition, specifying formats will help with consistency in the types of information resulting from spatial data.

9. Scientific Applications

In the FY06 project year, a key goal is to evaluate how the Toolkit relates to scientific efforts in critical ecosystem areas and how specific elements of the Toolkit might be amended to facilitate such work.

Identified Legacy Databases

Partners identified the following databases identified below in their priority order:

1. Series of Transects and Plots Documenting Plant Species

Title: Transect and Plot Legacy Database
Date of data collection*:* 1986-1988, 1990, 1992-1996
Overall Project objectives*:* This legacy database focuses on the establishment of a permanent monitoring system comprised of a series of transects and plots designed for the purpose of monitoring changes in vegetation across the grassy balds communities of Roan in response to long-term vegetation management. Information generated from the establishment of a permanent vegetation monitoring system (and from the baseline data on rare plant and animal populations) in the late 1980's was deemed essential to the future management of the Southern Appalachian balds, where there is a need to control natural succession to maintain high biological diversity and aesthetic

[3] Peine, John et. al. Bio-Science Evaluation for USGS NBII Applied to the Texas NAFTA Highway T&E Species Digitization Project, 7. 24 February 2005.

values. This legacy database encompasses many years of data collection and therefore is considered a multi-layered legacy database. In summary, this legacy database documents the following: (1) establishment of a permanent monitoring system, (2) two relocation efforts to re-monument the original transect and plot markers, (3) field data gathered at the transects and plots, (4) various analyses of the data to document baseline conditions and/or evaluate the effectiveness of management activities on the grassy balds vegetation composition and structure. Because this is a multi-layered legacy database, the objectives have been briefly summarized by metadata file:

a. Plot-Transect - 1986-1988 Installation
Principle Investigators**:** Jame Amoroso, Marj Boyer, Paul Hamel, Karin Heiman, Darlene Kucken, Laura Mansburg-Cotterman, Nora Murdock, Michael Schafale, Alan Smith, Paul Somers, Rob Sutter, Alan Weakley, and various Southern Appalachian Highlands Conservancy volunteers
Specific Project Objectives*:* The purpose of the 1986-1988 efforts was to (1) acquire funding to initiate active management of ecologically significant grassy balds communities on Roan, and (2) provide baseline data on current vegetation composition and rare plant and animal populations in the grassy balds communities of Roan in order to evaluate the effects of long-term management. The specific objectives were to:

- Obtain baseline data on the composition and structure of the vegetative cover types of the Roan balds to determine if future management (or natural events) is having the desired effect on the vegetation by (1) interpreting aerial photographs to determine the number of vegetation cover types present, (2) ground-truthing the aerial interpretation to determine accuracy and to ascertain dominant species of each type, (3) sampling vegetation at permanent transects and plots, and (4) photographing all transects and plots,
- Inventory the balds for the location and extent of eight rare plant species by (1) performing a systematic survey (via a walk through) and (2) mapping the rare plant occurrences on topographic map overlays,
- Obtain baseline data that will enable long-term monitoring of the effects of bald management on the rare plant species populations by (1) establishing permanent monitoring plots in populations of four of the eight rare species and (2) photographing all rare plant plots, and
- Conduct animal surveys to help determine sensitive or management indicator species in the balds communities by (1) implementing a variety of trapping and baiting techniques designed to maximize the diversity of mammal, amphibian, reptile and land snail species captured, and (2) performing point counts using variable circular plots to census birds.

Note that Metadata was compiled for only one of the four objectives listed in order to satisfy project purposes at this time.

b. Plot-Transect - 1987 and 1992 Mowing Analysis
Principle Investigators**:** Steve Simon
Specific Project Objectives**:** This component of the legacy database represents an analysis of the data collected from transects and plots on Round Bald sampled in 1987 and re-sampled in 1992. The specific objective was to:

- Examine the effects of four years of mowing on the grassy balds vegetation community.

c. Plot-Transect - 1990 Relocation
Principle Investigators: Karin Heiman and Darlene Kucken
Specific Project Objectives**:** This component of the legacy database contains information on a relocation effort conducted on all Roan grassy balds in 1990. The specific objective was to:

- Relocate and re-monument transects and plots on all grassy balds.

d. Plot-Transect - 1992-1993 Field Collections
Principle Investigators: Karin Heiman and Alan Smith
Specific Project Objectives: This component of the legacy database contains information on the 1992 and 1993 field collection efforts on various Roan grassy balds. The objectives were to:

- Re-sample vegetation at permanent plots and transects on Round and Jane Bald in 1992 to document the current grassy balds vegetation composition and structure in anticipation that changes in vegetation between 1987 and 1992 could be evaluated, irregardless of specific management activities, and
- Re-sample vegetation at a subset of permanent plots and transects on Round Bald, Big Yellow Mountain Bald, and Grassy Ridge Bald in 1993 to document the current grassy balds vegetation composition and structure in anticipation that changes in vegetation between 1987 and 1992 could be evaluated, irregardless of specific management activities.

e. Plot-Transect - Community and Mowing Analysis
Principle Investigators: Joe Noto, Steve Simon and Chris Ulrey
Specific Project Objectives: This component of the legacy data set contains information on an analysis of data collected at transects and plots in 1987-1988 and re-sampled in 1992. The specific objectives were to:

- Document the pre-management (baseline) conditions of the grassy balds vegetation (including rare plants) on all grassy balds in anticipation that response to long-term vegetation management could be evaluated, and
- Document the changes in grassy balds vegetation composition (including rare plants) on Round Bald and Jane Bald in response to the hand-mowing between 1987-1988 and 1992.

f. Plot-Transect – 1992-1994 Goat Grazing Study
Principle Investigators: Dave Danley, Gary Kauffman, David McFee and Steve Simon
Specific Project Objectives: This component of the legacy data set represents the data collected from a subset of transects and plots that were re-sampled in 1992, 1993, 1994 and 1995 to determine the changes in vegetation composition in response to active goat grazing (1992, 1993, and 1994) within the goat paddocks on Round Bald. The specific objective was to:

- Document the changes in grassy balds vegetation (including rare plants) in response to goat grazing.

g. Plot-Transect - Mowing and Grazing Analysis Summary
Principle Investigators: Steve Simon
Specific Project Objectives: This component of the legacy data set consists of a short summary of the monitoring results on Round Bald during 1987-1994. The objectives were to:

- Document a brief summary of the monitoring results from Round Bald, based on three separate analyses: (1) 1987 and 1992 line intercept data associated with mowing efforts, (2) 1987 and 1988 plant community data associated with baseline conditions, and (3) 1987, 1992, and 1994 plot data associated with the goat grazing study.

h. Plot-Transect - 1996 Relocation
Principle Investigators: Karin Heiman and Jame Amoroso
Specific Project Objectives: This component of the legacy database contains information on a relocation effort conducted on all grassy balds in 1996. The objectives were to:

- Relocate and re-monument transects and plots on all grassy balds using a GPS.

i. Plot-Transect - 1996 Relocation GIS
Principle Investigators: Tom Burley
Specific Project Objectives: This component of the legacy database contains information on the differentially corrected GPS data that was generated from processing of the original 1996 GPS files in 2005. The objective was to:

- Differentially correct the original 1996 GPS location data where possible.

2. Geo-Referenced Boundaries of Adaptive Management Activities (Mechanical Mowing and Hand Cutting)

Title: Bob Harvey's Briar Bashing Legacy Database
Date of data collection: 1989-1991, 1993-2002, 2005

Principle investigator: Bob Harvey
Project objectives: This legacy database represents information on Bob Harvey's annual Southern Appalachian Highlands Conservancy records that document the mowing that occurred on Round Bald and in Engine Gap during 1989 - 2005. The objective was to:

- Map the annual mowing efforts implemented by Bob Harvey's SAHC field crew and persons contracted by SAHC (D.R. mechanical mower and weed eaters, US Forest Service track mower) on Round Bald.
- Provide Bob Harvey and SAHC personnel with a very general portrayal of the vegetation response, particularly changes in black berry coverage, on areas mowed annually.

3. Herbicide and Mowing Treatment Plots

Title: Black Berry Eradication Study Legacy Database
Date of data collection**:** 1986-1987
Principle investigators**:** Joe Dabney, Dan Dunford, Tim Evans and Alan Smith
Project objectives: This legacy database documents data collected from a set of treatment plots established on Round Bald for the purpose of monitoring changes in grassy balds vegetation (particularly changes in black berry coverage) in response to herbicide application (Garlon 4 and Round-up) and mechanical mowing. The objective was to:

- Test the effects of mechanical mowing and herbicide application (Garlon 4 and Round-up) on grassy balds vegetation, and in particular study the possible solutions for controlling black berry growth.

4. GIS dataset produced under contract for The Nature Conservancy by Mr. Tom Kenney

Title: Tom Kenney's 1999 GIS Legacy Database
Date of data collection: 1999
Principle investigated: Tom Kenney
Project objectives**:** This legacy database documents GIS data that was compiled for the purpose of providing a GIS-based landscape tool for conservation planning on Roan at a landscape level. The objective was to:

- Compile GIS data on the physical and biological features of Roan to provide a landscape tool for conservation planning.

Application of Toolkit

Legacy databases selected to be applied to the Toolkit during FY05 include the following:

1. Transects and plots database.

The attempt to establish a baseline of vascular plant information has been extremely complicated in that data was collected approximately 10-20 years ago, was collected over a 10 year period, and involved at least 21 individuals on some level. In addition, the format of data collection has changed, the species identification codes have changed, the data sheets and reports are scattered around the Eastern U.S. and for the most part, attempts at re-location of the transects and plots have not been successful. Also, the extent of analysis of the data is not well known. According to one of the Draft US Forest Service Challenge Grant Proposal for 1986-1987, the plots and transects project concept was as follows:

> "To provide background data on current vegetation and rare plant and animal populations in the 'grassy balds' communities of Roan Mountain Massif in order to be able to evaluate effects of active management planned for 1987-88. Specific objectives listed in preferred order for funding are:
>
> 1. Mapping locations of eight rare plant species on Roan Mountain;
> 2. Ground-truthing USFS, aerial, infra-red (IR) images of the balds in order to determine and map current vegetation types;
> 3. Establishing permanent monitoring plots in populations of four of the eight species included above, and
> 4. Conducting animal surveys to help determine sensitive or management indicator species in the area.

> What is learned by the combined research and active management will provide information essential to management decision-making for other Southern Appalachian balds, where natural succession must be controlled to maintain high biological diversity and aesthetic values."

This database provides a poignant example of a range of obstacles that can occur when the principles associated in the Toolkit are not followed, and serves as a quantifiable example of the time, effort, and resources that might be wasted when a systematic process is not followed for project documentation and dataset handling.

2. Geo-Referenced Boundaries of Adaptive Management Activities (Mechanical Mowing and Hand Cutting)

The maps of adaptive management are a challenge in that the locations as identified on the maps do not coincide with latitude-longitude information included on the datasheet which is understandable due to the rapid increase in sophistication of hand-held GPS units and related satellite system. Nevertheless, these maps are an important documentation of the vegetation management that has occurred on Roan, going all the way back to 1989 in their documentation of Roan's management history. According to all of the partners, they are the only known existing documentation of these types of activities. Serving as a great conceptual model for more accurate procedures in the future, their conceptual idea combined with the specifications of the Toolkit will provide for reliable, accurate information for future adaptive management analysis.

Summary of the Example Priority Legacy Database Issues and the Value-Added by Utilization of the Data Management Toolkit

Overall, this first project year has shown that the two priority legacy databases – the Plot and Transect Legacy Database and Bob Harvey's Briar Bashing Legacy Database could have both benefited in several areas from the concepts presented in the Toolkit. Although application to the priority legacy databases is not possible after the fact, the Toolkit as applied to these example Roan legacy databases serves as a mechanism to evaluate the adequacy of data management and, consequently, the perceived utility and value of the data in that context. In turn, these evaluated examples can bring to light the current and future need for a systematic approach to data management.

1. Objectives, Planning, Compatibility

During the course of this first project year, it seems that specific long-term objectives related to evaluating the variable aspects of different types of vegetation management techniques (methodology, defined treatment boundaries, timing, intensity and frequency, etc.) were not sufficiently developed in projects so as to be able to provide specific recommendations for future adaptive management. Without these specific objectives, the specific types of data needed to meet these objectives could not be defined. This also brings in the issue of defined levels of accuracy, completeness, and other aspects of Quality Control that could not be defined without knowing the specific types of data that were required to meet the specific objectives for adaptive management. For example, important data related to mowing were not collected nor analyzed in terms of intensity and frequency, nor were mowed areas boundaries documented in a way that could provide for a confidence level of accuracy. It is worth noting that difficulty associated with documenting mowing can be partly attributed to inconsistencies in mowing uniformity due to differences in various vegetation densities, as well as due to physical features that create barriers for mowing equipment, etc. In addition, the lack of availability of georeferencing equipment such as GPS in addition to the less than desirable level of accuracy (at least during the late 1980's and 1990's) associated with such technology during the time is a factor (this is further addressed in Part 2 of this section).

The only known spatial documentation of mowed areas since 1989 has been by a SAHC volunteer, Bob Harvey, who created hand-drawn maps with approximate treatment (mow) boundaries on Round

Bald. However, this data did not include any aspects of frequency or intensity of the management, but simply were meant to provide a very approximate portrayal of the mowed areas each year. It is uncertain whether this data could have met any of the needs related to more specific analysis on mowing as described above. In either case, data of this nature were not incorporated into any analysis, in part because Roan partners who had done studies on mowing either were not aware of its existence, were uncertain of the utility of the data in this context, or simply had not defined objectives that would have called for such data to be included.

The broad objective of the initial plots and transects installation was to provide a community composition "snapshot" for the grassy balds in hopes of evaluating management practices. However, during the course of interviews for creating metadata for the legacy datasets associated with this project, a former Roan project partner noted that community composition is often used when management and monitoring objectives are not well defined, so a community composition study aims to simply collect as much data as possible. Due to this broad objective and the associated amount of data needed to meet this objective, as well as due to the large scale area of the balds, the less accurate method of ocular estimates was used so as to allow for the sampling to be performed quickly in order to cover the full extent of the balds. Because of this, it seems logical that at least two collectors would need to work on this data collection effort together, at least initially (which was done), and especially if data collectors change over the years. During interviews it was noted that this method can be subjective and that this could potentially be a significant source of variation in the data if field collectors did not closely collaborate on estimates during each sampling effort. In addition, if resampling in later years was to occur with lesser manpower and resources, the question of reliability and repeatability between years arises due to the potential subjectivity of this method. Central to this question is establishing precise and practical research question and articulate how it serves the needs of adaptive management and science.

An analysis of a subsample of the data early on (e.g., all of Round Bald data) before moving on to sample other balds as an effort to determine the minimum sample size needed may have proven beneficial by serving as a pilot for evaluating potential feasibility. However, due to vegetation types potentially varying significantly across balds, the assessment of minimum sample size may have been needed on a bald by bald basis. At the very least, a follow-up on the minimum sample size after the first field effort per bald could have proven beneficial for assessing the methodology used so as to ensure future feasibility.

Despite the methodology accommodations due to the large-scale area of Roan, the issue of long-term sustainability for such a huge effort remains an important factor as a result of the sampling intensity associated with the scale of the balds. Such an effort is not feasible today in the context of resources (time, money, manpower) needed to replicate the full initial effort. This, coupled with the inability to relocate a majority of the plots installed, has been a roadblock by significantly decreasing the value (and thus potential) of the original data collected. As a result, a full comparative analysis of changes in the vegetation communities of Roan and subsequent effects of management since the plot/transect installation has not been possible.

It is worth noting that such issues as budget cuts and decrease in manpower are typically not foreseeable issues that can certainly influence the repeatability of a project 5-10 years down the road. However, the critical issue of not being able to fully relocate the plots and transects prevents re-sampling for analysis of spatial-temporal change even if funds and man-power were available. Lack of good spatial data in any form is an over-arching issue with much of the legacy databases associated with Roan. This aspect is further addressed in Section B below. Because of this inability to relocate the original sampling sites, the question arises as to how much emphasis should truly be placed on some of the identified priority legacy databases in the context of planning for future work. In addition, even if they could be relocated, having to accept a lower level of accuracy and/or the less desirable of methodologies in order to make future work compatible with legacy data would essentially allow the data management issues identified in this report to persist. A dataset is only as accurate as the least accurate part of it.

In the context of the above issues, the Toolkit identifies key considerations in terms of long-range planning and the potential sustainability of a project in the context of data management. The critical issue of developing specific objectives and long-term objectives can help with planning for the types of data needed and how these needs will be met 5+ years down the road so that those objectives can be met. At the foundation of the Toolkit is the concept of data integration and all of the elements that must be considered, defined, and thoroughly documented so as to ensure a reliable, integrated, well-documented database geared towards those objectives beyond the initial work done. This approach can be thought of as a long-term investment approach by managing data to maximize its value both during and after the project for which it was collected. Strategic long-term organizational goals for data and information management can help ensure this. With large scale efforts such as the plots and transects, it can be easy to overlook certain aspects that end up having a much larger impact later on than originally anticipated. The Toolkit can help bring those aspects to light, hopefully allowing for a project and its associated data and information to be useful and reliable in the long term.

Compatibility and ability to integrate datasets are also issues that were identified associated with work done on Roan. These are aspects that the Toolkit emphasizes throughout a project via proper planning suggestions and the identification or establishment of standards in all aspects. This involves the critical need to be able to integrate various datasets such as those related to Roan in the context of adaptive management so that adaptive management effects on plant communities can be studied and evaluated. Several factors that were identified during this project which can hinder the ability to integrate are insufficient integrated overarching objectives, insufficient integrated data management planning, lack of communication among partners, lack of standard implementation procedures, lack of standard documentation, and lack of standard methodology associated with developing the datasets. One example of these factors mentioned was an instance involving an individual that had previously been involved with analysis done on plot data and specific management techniques. During a meeting for this project, Bob Harvey's Briar Bashing Legacy Database was mentioned, and it turned out that they were not aware that anyone had been creating maps showing mowed areas from year to year since 1989. This type of data was in fact an important aspect that was missing from analysis they had done on the effects of mowing. Though Bob's maps do not contain all of the missing puzzle pieces, they certainly serve as starting-point conceptual model for tracking this type of management in a spatial context.

For future work and analysis geared toward understanding the effects of different types of management on Roan's biological resources, factors influencing the balds (natural and man-made), specific management objectives for the balds, desired conditions, and subsequent required data to meet those objectives and conditions should be thoroughly analyzed and broken down so as to better understand how they effect each other. Development of a model based on factors such as these could allow for a better understanding of how they influence each other, and would allow for better analysis of these processes. By clearly defining the types of information needed to meet defined objectives, thematic data necessary to produce such information could be better identified and described. The development of any model is naturally an iterative process, but by applying a holistic, integrated approach, efficiency and organization can be realized. The Toolkit incorporates a conceptual, logical, and physical model approach overview as a suggested way for developing a more realistic and useful database pertaining to, in this case, adaptive vegetation management. In the long run, this would allow for better analysis and ultimately better information for driving decision making.

2. Spatial and Non-Spatial Data Elements

The importance of good spatial data cannot be emphasized enough. The only usable spatial data associated with these legacy databases were some Trimble GPS files from a 1996 plot and transect relocation attempt, as well as hand-drawn maps showing managed areas on the balds. The limited amount of spatial data was not captured with any quantifiable level of accuracy or completeness, nor seemed to have been done in any systematic or standard way. In cases such as this, it is difficult to rely on such data beyond a general portrayal of conditions or activities. This creates problems for compatibility with future work as well in that a dataset is only as accurate as the least accurate part.

In other words, if the data was viewed as relevant and usable for another project in terms of theme and representation, it may not meet the determined accuracy needs for that project. Much of the spatial data available existed in hardcopy format, and what little was in digital form was not in a GIS-compatible state until it was processed here at the University of Tennessee this past year.

Basic spatial data elements such as datum and projection, scale and data representation, as well as field methods and equipment parameters/settings were not recorded and seemed to not have been defined beforehand for the limited amount of spatial data available. Also, basic non-spatial data elements such as file naming conventions were not consistent nor defined, and limited documentation about what the data was intending to convey was often presented in a way that only the project participant could understand. In cases when known originators were contacted with questions, they were usually not able to recall the information needed due to the passing of time. The Toolkit presents these various factors in terms of spatial and non-spatial data elements and their importance in the context of project planning and Quality Assurance. By being aware and defining these important aspects of data at the outset, a project's ability to meet expectations and objectives will be greatly enhanced.

It is worth noting that many of the legacy databases that were identified as priority (primarily those associated with the plots and transects before the 1996 relocation attempt) occurred when GPS utility and technology was not as widely available as today, thusly it is understandable why they may not have geospatial data associated with them. In speaking with one contact for the Cherokee National Forest during the metadata creation process, they stated that a handful of GPS units were acquired in 1992, but for the most part GPS usage did become common until Selective Availability (SA) was turned off in May of 2000. Selective Availability was a policy adopted by the Department of Defense that intentionally degraded the quality of GPS signals and accuracy levels obtainable by non-military GPS users. Selective Availability, in addition to the mountainous terrain of the Appalachians, was an issue with any GPS usage during the 1990's in that getting a signal could be difficult, and due to SA the accuracy could be very questionable. However, the balds are some what of an exception in terms of surrounding topography due to the lack of canopy and the fact that they have a higher elevation located on mountaintops.

With increased usage of GPS after 2000, the use of recreational grade units as opposed to mapping and/or natural resource grade units has been an issue in the context of the spatial data elements identified herein. The Cherokee NF contact stated that there is often a problematic misconception by people that all GPS and related technology is equal in terms of accuracy and other spatial elements due to using the same satellites and similar misconceptions. In addition, frustration can be experienced when attempting to use a natural resource grade unit with parameters that filter satellite signals for accuracy purposes, particularly in mountainous areas. One viable option, as noted in the Toolkit, is that mission planning using GPS software should be considered so as to determine the optimum time for GPS signals at a given location.

The issue here is that recreational units do not filter signals, so the accuracy levels obtained are often uncertain, and in fact may be no better than simply drawing features on a quad map of the area. However, recreational grade units will be able to gather data and receive signals where more accurate resource grade units might have difficulty. Again, on the balds this should be lesser of an issue. In the context of projects and project needs, this essentially presents a question of quantity or quality, a crucial element that should ultimately be addressed and decided on at the outset of a project via a Quality Assurance plan (discussed below). In the case of having more spatial data as opposed to good spatial data, however, the benefits of the former are unknown.

In another conversation during the fall of 2005 with another Forest Service employee, as well as with a local Knoxville GPS specialist that had previously interned with the Forest Service, both noted that issues with spatial data such as these identified in this report are still very much prevalent with projects and work done today in spite of widely available equipment and expertise. The Cherokee NF contact had noted that there really is not a top-down management mandate for the gathering of good spatial data with specified spatial data parameters in the Forest Service. Also, all three noted that these problems persist today due to many people either not being comfortable with the technology and/or not being cognizant of the need for good spatial data. This is a particularly important issue and problem because approximately 80% of all data has a geographic or spatial component. As a result of these problems, projects end up generating data with limited utility, comparability, and compatibility. And, as the current USFS employee noted in frustration, projects and data produced with limited utility are simply a waste of resources (time, funds etc.)

The Toolkit emphasizes the relevance of spatial data by identifying critical spatial data elements so as to incorporate those into the project planning process. It also provides suggested guidelines for spatial data acquisition and processing with GPS technology that is non-brand specific, as well as for digitally enabling hardcopy spatial data. As with any project, these elements should all be defined at the outset so as to ensure that the objectives of the project are met. The Toolkit's presentation of these concepts and suggested guidelines can serve as a starting point for things to consider. In turn, consistency of methods and data elements such as these can also help promote compatibility among datasets through a consistent level of detail. With the advances in GPS technology, as well as available expertise and equipment, lack of good spatial data should not be an issue today.

3. Consistency, Quality Assurance and Quality Control

Reports and analysis done on vegetation management activities using available data typically could not make specific recommendations for application in terms of intensity, frequency, timing etc. Specific data needed to fully evaluate vegetation response to a type of management were often not collected and/or not incorporated into the analysis. Data analysis was often not done consistently, and in multiple instances (usually not documented as to why) only a portion of a dataset would be subject to statistical analysis. In other cases, data directly related to objectives for the project, for unknown and undocumented reasons, were not used at all. When known data and project originators were asked about these cases during interviews, typically they could not recall why this was the case. Other aspects including undefined/undocumented data elements such as field definitions and attribute types/specifics (such as species codes) were noted by data originators as being problematic. Some originators that were interviewed noted that different types of species codes were often used during different years, resulting in considerable time and effort attempting to cross-walk data if any comparative analysis was to be done. In addition, the originators had poor recollection of the specifics regarding species codes now 10-20 years later. The Toolkit thoroughly addresses issues such as these by advocating the use of standard documentation or metadata, none of which was available for any of the priority legacy databases prior to this project.

Based on documents and information available, it seemed that Quality Assurance issues such as accuracy, completeness, and consistency requirements in all aspects of projects were not well defined for much of the work related to Roan. Underlying this is unquantifiable error likely resulting from the primary sources of error in a project: the source data, data entry, and data analysis. It is imperative that the level of detail be consistent throughout a project. Without established and adhered to standards for these aspects of a project at the outset, as well as defined procedures for Quality Control, the integrity of data and resulting information will likely be compromised, ultimately creating a lack of confidence in the resulting information. When that happens, the value of the data is for all intents and purposes, gone. In turn, this may result in the inability to fully satisfy project objectives. The Toolkit addresses the important aspect of Quality Assurance and Quality Control with an overview of QA concepts that are largely geared towards spatial data, but conceptually can be applied towards other aspects of a project. These concepts include defining data acquisition methods, acceptance criteria and sampling strategies, and parameters for analysis. The Toolkit also provides some suggested Quality Control methods utilizing common GIS software. As well, the Toolkit's conceptual presentation of standardized methods for all aspects of a project in conjunction with proper documentation and the defining of data expectations all contribute to Quality Assurance and Quality Control from the outset. This, in turn, will help promote compatibility and utility among datasets created with similar standard parameters.

4. Standard Documentation

Lack of standard documentation was one of the biggest issues faced during this project. In order to create FGDC metadata (to the extent possible/feasible), extensive research was required simply to determine who the data originators were and their relation to a project as that information was usually incomplete or not readily available at all. There was uncertainty surrounding many project elements, and where documentation did exist, it was not consistent or standard between years. During the research process, it became evident that a majority of Roan partners were not aware of all work related to Roan over the years due to not having standard metadata documentation available. One poignant example of inefficiency resulting from unavailable standard documentation involved different people doing the same time-consuming work with the same data. In this instance, a portion or all of the same data from the plots and transects dataset had been entered electronically four separate times over the course of about five years. Another example worth noting is that two compositional analyses were conducted using the same data, and that the data analysts conducting the second analysis were not aware that the first analysis had been done.

In some cases where documentation did exist, it was only available in the form of internal agency documents, as well a notable number of documents and reports were still in draft form. Locations of hardcopy data, documents, and reports were scattered among home residences (home offices and storage units) as far north as Massachusetts and as far south as Mississippi, as well as among several agency and organization offices. Typically, documents in such locations were not organized in a systematic way and were often incomplete. As well, important details about some of the limited spatial data associated with priority legacy work (Bob Harvey's maps) on Roan were determined lost due to metadata having never been created, in addition to the associated original contributing digital data being lost due to what was believed to be equipment changes (swapping out of computers, etc.).

Some information could only be discovered after talking with known data originators, however most are in different jobs now, some many states away (as far north as Massachusetts and as far south as Mississippi). Because of this, understandably, time constraints on their part created difficulty in getting needed information from them, and often time individuals had to be contacted several times. Face to face interviews proved to be the best way for getting in formation, but these required travel time on the part of the researcher and in the case of any interviews, time was invested by both the interviewer and interviewee. Even then, originators that could be reached often could not remember important details, and on several occasions information generated from interviews contradicted information presented in data documents or gained from other data originators. Where discrepancies existed, generalizations had to be made. Also, some originators noted that they had retained documents in hopes of finalizing analysis or reports, or to simply satisfy concerns that data and information might be lost within the state and federal agency domain.

Because of the limited, non-standard documentation and uncertainty of information gained from originators due to passing of time, many unknowns remain despite having created FGDC-compliant metadata to the extent practicable for these two legacy databases. One instance is uncertainty about what constituted a 'full' or complete set of photographs related to the plots and transects. Some of the others have been addressed in latter parts of this report including: species codes, explanation of content in electronically entered data, field methodologies, analysis procedures, QA/QC procedures, and accuracy and completeness of data to name a few. As well, the extent of the metadata documentation that we compiled in terms of completeness 10+ years later can only be approximated. It should be noted that Executive Order 12906, established in 1994, in fact mandated the creation of metadata for datasets from January 1995 forward for all Federal Agencies, and also stated that those agencies are responsible for devising plans to document data previously collected or produced (legacy) to the extent practicable. Some of the legacy databases addressed with this project and some of the analysis and work done occurred during 1994 or prior, so a mention of hindsight is appropriate here. However, plans for metadata creation for those legacy databases were part of the mandate but unfortunately in most cases were not carried out. As demonstrated by this project, the creation of metadata this far after the fact, despite a Federal mandate, has proven extremely laborious and inefficient due to having to locate, understand, and then compile relevant information 10+ years later.

The Toolkit addresses the need for standardized documentation by including a section on the FGDC-NBII biological profile extension metadata standard. The Toolkit integrates the idea of developing standard metadata documentation throughout the life of a project and/or dataset as opposed to leaving it as an afterthought that may or may not get done once a project is over. Included are mention of free software programs for metadata creation, as well as information on issues associated with metadata, a questionnaire-style form for extracting relevant information with a crosswalk to the appropriate FGDC-NBII standard sections, and suggested workflows for creating an FGDC compliant metadata record using the free tools and resources described in the Toolkit. As well, the mention of metadata clearinghouses is included so as to make the existence of datasets and information about them "discoverable" by others. This, in turn, can help reduce the occurrence of double-work as demonstrated by the example above, and can also help promote leveraging of existing resources, better communication, and increased collaboration among heterogeneous agencies with common goals on the same landscape.

5. Data Access and Archiving

The disparity of data, document, and report locations related to Roan over nearly 20 years, besides the lack of standard consistent documentation, was one of the more problematic areas identified during this project. In general, access to original materials and proper archiving/custodianship was nonexistent. Extensive and time-consuming research had to be performed to locate people who knew about locations of documents, and then time had to be invested by those individuals to uncover those materials and in order to provide them to us. A process like this proved time consuming for both the researcher and the individual in possession as documents and documents were usually exchanged during face to face meetings. This was due to the concerns of many regarding the shipping of original and only existing versions of datasheets, documents and reports in that losing them in the mail system could be detrimental. In cases where only draft versions of materials existed, the most 'current' draft version had to be determined and sometimes assumed.

Materials related to Roan were scattered all around the Southern Appalachian region, as well as far north as Massachusetts and as far south as Mississippi. Locations included home residences/personal offices, storage units, as well as multiple agency offices. As previously stated, many had kept original data and documents in hopes of eventually finishing reports and/or analysis, though the fact that many of these people were in different jobs now made this highly unlikely. Also as previously mentioned, many were concerned that materials would be lost in the state and federal agency domain, thus clearly suggesting that scenarios such as this are not unique and that people are aware of these data management inadequacies in multi-agency collaborative efforts. An example of the effects of this on completeness and accessibility of materials is the original photograph set associated with the plots and transects installation. Documentation uncovered during the search process for this project noted that an inventory done in 1996 showed that some were missing, some were without labels, and some were too obscure due to exposure problems. A similar scenario plagues datasheets associated with these projects. During the search process it was determined that multiple originators had original copies, and occasionally versions would turn up attached to other documents and reports as a reference for that report.

Other examples of effects of this include instances previously mentioned where various researchers manually performed data entry and/or other work related to Roan several times simply because they had no way of knowing that it existed. As further proof of this problematic issue from an outside source, a study done in 2001 estimated that about 50% of the federal government's geospatial data at the time was redundant.[4] A centralized archive related to Roan materials in conjunction with an established data policy and proper metadata documentation could let researchers know what exists and where to find it. This would in turn allow for the leveraging of existing resources if an identified

[4] United States. General Accounting Office. Geographic Information Systems: Challenges to Effective Data Sharing. Testimony of Linda D Koontz, Director of Information Management Issues before the Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census, Committee on Government Reform, House of Representatives. 10 June 2003.

dataset could meet some or all of a project's data needs, thus helping to avoid costly replication of work.

After several months of research, many documents are still geographically scattered, and to what extent 'complete' sets of some exist is unknown. This issue also relates and contributes directly to the lack of standard and consistent documentation in that proper documentation should have been developed during the full lifecycle of these projects. Individuals now typically do not have the time and resources needed to dedicate to such an extensive search as has been performed through this project. Without the needed materials and information, metadata cannot be created, thus creating a problematic cycle. People move on to other jobs, memory fades due to time, and such occurrences as equipment failure and/or changes result in actual materials as well as information associated with the locations of those materials being lost in the shuffle.

The Toolkit can assist in developing an appropriate data policy regarding storage of data and information, access, archiving, and custodianship by covering these areas, what they entail, and then by providing specific suggestions and recommendations where possible. The Toolkit suggests that these important issues be fully addressed at the outset of a project via the Quality Assurance plan and associated project scoping documents. This helps ensure that a project is properly prepared to manage the resulting data and information resulting from it, creating the idea of a full project life-cycle from project inception to giving the data a properly managed "resting place".

A great example of how to address this has been re-initiated as a result of this project. The Archives of Appalachia at ETSU has agreed to house and manage materials related to Roan in order to establish a fully functional Roan Archive with proper storage conditions, access, and professional custodianship. A key feature is the potential to restrict access as needed. This great example of institutional buy-in and leveraging of resources demonstrates that existing expertise, often times, can be taken advantage of and applied to significant natural resource efforts. In addition, by being cognizant of and by proactively addressing these issues from the outset of a project, the important long-term investment approach is integrated into the project which helps to ensure the value and utility of data and information down the road for future work.

When efforts and the people contributing to them span across several agencies, a common data management framework must be established so that common goals on the same unique landscape can be achieved. All aspects of a project must be assessed in the context feasibility, as well as sustainability if the project is to have a long term scope. Specific objectives, specific data requirements, and defined standard methodologies should be established during the planning process in order to identify data needs, promote consistency, and to ensure compatibility of the data and information with future efforts. To help ensure that these methodologies are carried out, a Quality Assurance plan and Quality Control procedures should be in place and integrated with the latter project elements. Underpinning all of this is continuous communication between all parties in order to hedge against misconceptions, as well to make sure that all key areas have been addressed from all relevant perspectives. Continuous dialogue is crucial for establishing an agreed upon, systematic methodology focused on similar objectives with emphasis on compatibility between present and future datasets if needed and where possible. The Toolkit establishes this idea of data integration via thorough planning by cross-walking data aspects back to a project's objectives, the project and goal requirements in relation to standardized methodology, aspects of spatial data and non-spatial data elements, proper documentation, QA/QC, and general data management needs as applied to the full project life cycle, ultimately helping to establish the view of data as being a long-term investment.

Detailed Evaluation of Example Priority Legacy Databases via Application of the Toolkit

The following important aspects of any project and consequently he Toolkit are applied as components of the evaluation, in detail, for the two case examples.

1. Identification of Those Involved with the Creation of the Database and Their Roles

a. Transect and Plot Legacy Database
Information on data originators, their respective roles, and their current contact information was not readily available in data documents, and where available on some level it was incomplete. Because of poor documentation, uncertainty developed as to who was involved with what part of the project, in addition to when and how the individual was affiliated with a particular project. In some instances, there was overlap between agency field collection efforts, data entry and data analyses. In some cases there was no information available in data documents on particular efforts and this information came to light only after talking with fellow data originators. As a result, an extensive search was required to locate this information and thus gain an understanding of the data and its utility thereafter. The following steps were implemented: (1) existing data documents were extensively searched for project participants and their respective roles, (2) a web-based query was performed to locate current contact information, and/or (3) requests for information on fellow project participants were made during interviews with the various data originators when needed.

The majority of current contact information was compiled through web-based queries because nearly all data originators have dispersed to other jobs located within the vicinity of Roan and/or many states away. Most information on individual roles came from interviews, and particularly interviews with data originators who have remained in the nearby geographic vicinity of Roan through the years and maintained a general interest in following the management on Roan. In most cases, defining individual roles was a gradual process; information was updated as it became available.

b. Bob Harvey's Briar Bashing Legacy Database
Identification information was acquired from the Southern Appalachian Highlands Conservancy whom Bob Harvey voluntarily collected the data for because information was not readily available in the data documents.

c. Black Berry Eradication Study Legacy Database
Identification information was available for only one out of four project participants. The one data originator interviewed did not have information on the other participants and had minimal knowledge of their affiliation with the project due to turnover in personnel between the 1986 and 1987 field seasons. Although a web-based query was performed and requests were made from various agency personnel who were most likely to have knowledge of past project participants, information for the latter three data originators and their roles, where available, came solely from data documents.

d. Tom Kenney's 1999 GIS Legacy Database
Identification information was obtained from a web-based query and existing data documents.

2. Clarity of Statements of Project Objectives and Intended Applications

a. Transect and Plot Dataset

- **Plot-Transect - 1986-1988 Installation**
  Aerials
  Aerial photos were never used to determine the number of vegetation cover types, nor to ascertain dominant species. Data originators do not recall exactly why aerial photos were proposed as a tool in this context, and know that aerials were used to map the rare plant populations on Roan as proposed originally.

  Community Classification and Stake Installation
  Vegetation was sampled at 897 plots and along associated transects installed across the balds complex in 1987 and 1988 to document the baseline data on the composition and structure of the balds vegetation in an effort to determine if future management is having the desired effect on the vegetation.

The 1987 and 1988 data was subject to rigorous exploratory data analysis using the Statistical Analysis System (SAS). Data was analyzed using Cluster Analysis Techniques, Discriminate Analysis procedures, and Principle Component Analysis to classify vegetation communities for all balds. Analysis included a comparison of the relationship between all field data collected in 1987 and 1988: line intercept, plot, and physical parameters.

Two major inventories were generated: (1) The first vegetative community composition inventory was compiled. Six vegetation communities were identified across the balds. These pre-management (baseline) conditions could be used to measure changes in balds vegetation community composition in response to future management activities (or natural processes), and (2) The most complete plant list, including rare plants, was compiled. A total of 197 taxa were identified.

There seems to have been some amount of oversight in understanding the feasibility of using this study design in the future in terms of projecting the likelihood of available funding and man-power at a later date. Given the mere fact that a series of permanent transects and plot markers were installed, there was some level of consideration for assessing the effects of future management by data originators during the development of the study design in the late 1980's. However, in the late 1980's when the field design was being developed and data was being collected for the first time, funding was not as limited a resource as it increasingly has become up to the current time. Therefore the feasibility of conducting a comprehensive and thus comparative re-sampling on the same scale as was performed in the 1980's has not been possible at later dates due to the greater limitations on funding and man-power. Instead, it appears that evaluations of management practices post 1987 and 1988 field collections, has taken place on various balds (notably Round Bald) and on a periodic basis when funding has became available and interests peaked.

Furthermore, at the time that the study design was being developed, no specific objectives were outlined in terms of identifying specific management treatments. In other words, no specific groundwork was drawn up a priori to identify specific treatment types, areas, methodology, and timing on which these treatments would be implemented in order to facilitate systematic documentation of vegetation response to management activities and thus ensure that the initial effort was worthwhile years later.

In terms of analysis, it should be noted that vegetation community classification analysis is a very subjective analysis. In other words, the analysis is influenced by whether the data analyst tends to lump or split. Data originators have noted that data analysis needs to be conducted again at this later date using more current software.

Photo-documentation
Photographs were collected to document changes in vegetation conditions and relocate permanent markers in the future. It is unknown exactly how many original photos existed as a result of the 1987 and 1988 efforts because this is not well documented in data documents. However, it is known that by 1996 there was an incomplete set of photographs.

- **Plot-Transect - 1987 and 1992 Mowing Analysis**
  Analysis was conducted on a subset of the line intercept data (dominate species) collected on Round Bald in 1987 and 1992 by calculating the changes in species dominance between pre-maintenance (1987) and post-maintenance (1992). Only a subset of line intercept data was able to be analyzed because only a 1957 meter section of the transect line could be accurately compared between years due to other management activities in nearby areas and the coverage of the mowed areas.

  Based on the overall results from the analysis mowing has been shown to reduce woody growth and may improve conditions for protected, threatened, and endangered plants.

Information on mow frequency and intensity were not incorporated into the analysis because the information had not been compiled by the time that the analysis was conducted. Data originators noted that mowing was irregular and that it is possible that a small proportion of the line intercept data analyzed might have fallen outside the confines of the mowed areas. Therefore some vegetation along transects would not have been exposed to mowing and therefore vegetation would not have been responding to mowing. Mow frequency and intensity, and careful documentation of mowing boundaries in general would have been needed in order to make a full interpretation of how mowing had an effect on the vegetation between 1987 and 1992.

In addition, a relatively small amount of data was analyzed relative to the amount of area that was mowed annually on the balds complex, which indicates that there was oversight in the amount of pre-planning that went into designing a study that tested the effects of mowing.

- **Plot-Transect - 1990 Relocation**

  Data originators consulted existing field notes and transect and plot photos during field efforts. An attempt was made to revisit all plots and an inventory was compiled for transects and plots relocated, re-staked, and re-monumented in the field.

  Relocation attempts were not as complete as they could have been because of limited funding and the inherent difficulty that field researchers had with relocating some plots and transects. If the plots or transects were very hard to locate, the field collector moved on to the next plot and transect so that as many as possible could be relocated in the amount of time allotted. Field collectors noted that additional funds would have made the relocation effort complete. In particular, no relocation was attempted for the following areas:
  (1) Entire Bradley Gap (approximately 5 transects),
  (2) Hump Mountain Bald - transects A, B, C, D, R-1, and R-2 (approximately 6 transects), and
  (3) Entire slope on Little Hump Mountain Bald (approximately 6 transects).

  Although these are permanent plots and transect markers, a significant number (less than half of the original number installed) could not be relocated. Field researchers identified the following as possible sources of stake removal, absence, or oversight: damage by mowers, misplaced or removed due to weather (freeze-thaw cycle) or recreationists, underlying physical features or high recreational use never allowed for installation, or encroachment of woody vegetation or soil deposition has resulted in oversight. Based on the low return rate of these permanent markers, it appears that this permanent marking design and the technique for relocating the markers were not highly effective 3- 4 years after installation on the Roan grassy balds.

  Furthermore, the transect lines have been difficult to relocate because compass declinations were not consistent among multiple field researchers nor were bearings always consistent for all transect lines, even though initial attempts were made to control this aspect of the design. It was difficult to keep the measuring tape straight in the windy conditions characteristic of the open balds, maintain the same bearing due to the natural arching topography of the balds in association with long transect lines in some areas, and standardize the compass declination given that multiple field personnel participated.

- **Plot-Transect - 1992-1993 Field Collections**

  In 1992, vegetation was re-sampled at permanent plots and transects on Round Bald and Jane Bald to document the current grassy balds vegetation composition and structure in anticipation that changes in species composition could be evaluated by comparing the 1987 and 1992 data. This data was not collected with the intention that specific management activities would be evaluated, but instead to document the changes in response to natural succession, irregardless of active management. In 1993, vegetation was re-sampled at a subset of permanent plots and transects on Round Bald, Big Yellow Mountain Bald, and Grassy Ridge Bald to document the current grassy balds vegetation composition and structure in anticipation that changes in species composition could be evaluated by comparing the 1987, 1992 and 1993 data. Again, this data was not collected with the intention that specific management activities would be evaluated, but instead to document the changes in response to natural succession, irregardless of active management.

In 1992 and 1993, no specific plans were identified for analyzing the data to determine if there were changes in vegetation composition and structure prior to re-sampling. Instead, at a later date, the 1987 and 1992 data were compared to determine the changes in plant composition and structure of the grassy balds vegetation between 1987 and 1992 (See USFS Roan Community and Mowing Analysis). The 1993 data was never analyzed.

Furthermore, although data collected during 1992 and 1993 were not collected for the purposes of evaluated the effects of management tools, note that if managed and unmanaged areas could be distinguished and there was a particular interest among project partners to evaluate the effects of a particular management tool, such as mowing (see USFS Roan 1987 and 1992 Mowing, USFS Roan Community and Mowing Analysis, and USFS Roan Mowing and Grazing Analysis Summary) or goat grazing (see USFS Roan Goat Grazing Study and USFS Roan Mowing and Grazing Analysis Summary), then analyses were conducted.

Overall, a relatively small proportion of the data collected in 1992 was used by various data originators at later dates to evaluate the effects of management (mowing and goat grazing) on grassy balds vegetation, and again none of the 1993 data was analyzed. In other words, no objectives were outlined in terms of identifying specific management treatments prior to data collection, but instead data was analyzed where management happened to correspond with data collected.

- **Plot-Transect - Community and Mowing Analysis**
  The 1987 and 1988 plot data (n=899) for all grassy balds was analyzed using TWINSPAN and Discriminate Analysis to classify the vegetation communities across all balds to document pre-management (baseline) conditions for the purpose of evaluating management activities post 1987 and 1988. In order to evaluate the changes in vegetation composition on the grassy balds in response to the hand-mowing between 1987 and 1992 a subset of the plot data (n=253) collected on Round Bald and Jane Bald in 1987 and 1992 was analyzed using Discriminate Analysis.

  The overall results from the community analysis using the 1987 and 1988 field data showed that there were 12 different community types across the balds. Based on the overall results from the analysis used to determine the effects of mowing, there was a shift in community types between 1987 and 1992. However, the shift in community types is assumed to be in response to mowing. However, the shift in community type may be occurring in response to natural succession. Detailed records on mowing intensity and frequency are needed in order to substantiate the assumption. The effects of mowing on protected, endangered, and threatened plant species on Round and Jane Bald will need to be gleaned from data documents.

  Only a small portion of the data entered in 1994 was analyzed to determine the effects of management (notably the effects of mowing) relative to the amount of data that had been compiled. [It appears that Simon submitted reports (goat grazing and mowing reports in 1993 while the summary report was submitted in July 1994)]

  Mowing frequencies and intensities were not included in the analysis because there was not a full understanding of the mowing frequency and intensity in the area, and therefore there was not a full understanding of the vegetation response upon which to evaluate the effects of mowing. Also, an assumption was made that all areas re-sampled in 1992 were exposed to mowing because it was sometimes difficult to determine where mowing had and had not occurred. In other words, mowing was not uniform in all places.

Data for physical parameters and line intercept data was not included in the analysis because there was interest only in analyzing the plot data.

Data originators noted that the general computer and software technology has changed since just 1994, and as a result noted that although the interpretation through data analysis was sound, they recommended that newer software be used to analyze the data at this point in time.

- **Plot-Transect – 1992-1994 Goat Grazing Study**
  Goats grazed on Round Bald during 1992, 1993 and 1994. Data collected in early 1992 provided baseline data prior to goat grazing. The line intercept data (primary and secondary dominant species) along 260 meters of transect within the confines of the goat paddocks in 1992 and 1995 were compared using basic descriptive statistics (percent change). Plot data (species cover and species composition) from 25 plots within the confines of the goat paddocks was compared among 1992, 1994 and 1995. This subset of data was compared to determine changes in percent species cover and composition between 1992 and 1994, and 1992, 1994, and 1995.

  Based on the overall results from the goat grazing study, goat grazing (1) has been shown to reduce woody vegetation, (2) encouraged grasses, sedges and herbs, and (3) may have improved the conditions for protected, endangered and threatened plants, notably Gray's Lily. Three years after the goat grazing had ceased the blackberries were back to their previous densities prior to goat grazing. A cost analysis of the study was compiled and presented to potential permittees through interviews conducted by the U.S. Forest Service. However, potential permittees determined that the profit margin was too low, given that the permittee would have to pay for expenses associated with establishing and maintaining fencing and watering systems and interests in goat grazing was minimal in general because it has become a small localized hobby in recent years. However, to fully understand the thoroughness of the cost analysis U.S. Forest Service personnel should be interviewed for a second time and associated U.S. Forest Service files need to be obtained if they exist.

  A full understanding of the goat grazing study details 12-13 years later is difficult to determine due to incomplete documentation and discrepancies in the data documents and interviews. Therefore, whether or not the original objective was completely met is unknown. Field methodology and field collector information between all years is incomplete. In one instance, conclusions were made based on the effects of mowing even though mowing was not mentioned elsewhere in the corresponding data document. It is has not been possible to document with certainty from which field effort this data was compiled and analyzed. The data most likely came from two separate sources: (1) field efforts conducted during 1992-1995, and (2) field efforts conducted in 1992 and 1993.

  The amount of data analyzed relative to the amount of data that was initially collected, which would have been represented by the area grazed by goats over the years, is unknown. It appears that only a small portion of the data that would have been collected was analyzed. Data that was supposedly collected in 1993 was not included in the analysis. There is concern that statistical analysis of data was conducted for some years and not others which causes concern for the ability to compare data among years.

- **Plot-Transect - Mowing and Grazing Analysis Summary**
  A brief summary of the monitoring results, based on three separate analyses, using a data from Round Bald was reported: (1) basic descriptive statistics using a subset of the 1987-1992 line intercept data was used to document the effects of mowing; (2) statistical analysis of the 1987-1988 baseline data was used to document the balds vegetation communities prior to active management; and (3) statistical analysis using a subset of the 1987-1994 plot data was used to document the effects of goat grazing.

Goat grazing and mowing have been shown to reduce woody growth and may have improved conditions for protected, threatened and endangered plant species. Herbaceous species (mostly exotics) tended to increase in response to grazing then mowing probably due to increased soil disturbance. The three methods used to analyze change in vegetation on Round Bald (line intercept [species dominance], plant community [community and species cover and abundance], analysis of species cover on plots) tends to result in the same conclusions.

- **Plot-Transect - 1996 Relocation**
  Data originators consulted existing field notes and transect and plot photos during field efforts. Plots were re-staked when located and an inventory was compiled on transects and plots relocated and re-monumented in the field using a GPS.

  Approximately 16 transects and 33 plots were relocated. Of the 33 plots relocated only 19 were re-staked. Some plots relocated were not re-staked or no attempt was made to relocate some transects and plots (R1 and R2 on Hump Mountain). In addition, some GPS files collected in the field were noticeably lost during field collection efforts due to unit malfunction (all plots on Grassy Ridge and plots along transects 1 and 2 of Round Bald). Again, because few permanent markers were able to be relocated, it appears that this permanent marking design and/or the technique for relocating the markers were insufficient for relocating the markers 9-10 years after installation. In addition, the attempt to re-monument permanent markers using a GPS was insufficient.

- **Plot-Transect - 1996 Relocation GIS**
  The original Trimble Geo-Explorer III files provided were differentially corrected using Trimble Pathfinder Office software, where possible, with base station files obtained from the NCGS. However, due to differential correction being done nine years later, not all of the base station files were obtained. For files that had no base station file obtained for them, the data points were simply grouped into one waypoint.

  During the processing of the original files it was determined that no GPS coordinates were available for transects and plots on some balds (Hump Mountain Bald and Bradley Gap). Nine files containing geo-referenced plots and transects located on Little Hump Mountain and four files containing geo-referenced plot and transect locations on Yellow Mountain were able to be located using NCGS base station records.

b. Bob Harvey's Briar Bashing Legacy Database
Annual maps of the mowing efforts were generated by Bob Harvey, a SAHC volunteer, to document the general effects of the mowing (particularly the changes in blackberry cover) on the grassy balds vegetation. On the ground mowing efforts documented were implemented by Bob Harvey's SAHC field crew (weed eater primarily; lopping and D.R. mower secondarily), and persons contracted by SAHC (track mower), and US Forest Service personnel (weed eaters and track mower) on Round Bald and in Engine Gap.

No descriptive or statistical analysis was conducted. Documentation of mowed areas was not collected in a systematic manner among years or in a manner that allowed for statistical analysis. Mowed areas were either sketched in their entirety (1988), sketched onto a U.S. Forest Service basemaps (1989-1992, 1994-2000), or sketched onto a U.S. Forest Service basemap in addition to GPS mow boundaries (2001-2005). Coverage of mowing by the above mentioned crews was not compiled in 1993 because Bob Harvey did not field his own briar cutting crew this particular year. The maps present a general portrayal of where the mowing has been implemented over the years on Round Bald and in Engine Gap.

c. Black Berry Eradication Study Legacy Database
Basic descriptive statistics using data collected in 1986 and 1987 from three types of treatment plots (Garlon 4, Round-up and mechanical mowing) on Round Bald was used to test the effects of

mechanical mowing and herbicide application on grassy balds vegetation, and in particular blackberry growth.

Based on the overall descriptive analysis, both types of herbicide treatments reduced blackberry growth by killing a notably large amount of the blackberry but would not control blackberry in the long-term. However, herbicide treatments also killed non-woody vegetation such as grasses, sedges and moss, which are desired vegetation. Mowing does not have a long lasting effect on blackberry (at least the way it was done and at the time of year that it was applied).

At least one data originator noted that herbicide application was not applied uniformly across all treatment plots and there appeared to be spillover into adjacent plots.

d. Tom Kenney's 1999 GIS Legacy Database

Existing GIS data was combined with derived GIS data that Tom Kenney developed. Existing GIS data consisted of spatial data on elevation, landcover, rivers and streams, political and administrative boundaries, roads, and biological/natural features. Derived GIS data consisted of spatial data on broad vegetation (landscape) zones, high elevation natural community areas, known habitat areas for significant animal species, and particular landscape site components. In addition, land ownership data for areas of three counties within the landscape were acquired or digitized.

2. Locating the Datasets

a. Transect and Plot Dataset

A total of 15 out of 21 people contacted were able to supply information which ranged from providing a few valuable comments to physically supplying data documents. Data documents were located at nine different physical addresses. Data documents were located in files of the following: (1) home residences (to include home office and storage units), and (2) agency and organization offices (SAHC-Asheville, SAHC-Kingsport, USFS-Burnsville, USFS-Asheville, Forest Service, USFWS-Asheville, NCNHP-Raleigh, etc.). More specifically, many documents retrieved from agencies and organizations, had been stored in "storage" file cabinets. That is, documents were retrieved from storage units of the NCNHP-Raleigh, USFS-Burnsville, USFS-Asheville, and SAHC-Kingsport offices. Data documents were scattered locally in Tennessee and North Carolina to as far ranging geographic areas as Massachusetts and Mississippi. Data documents have been scattered as a result of multiple changes in agency and organization personnel over the years.

Some field researchers retained the data in their personal possession to finish analyses or in an effort to satisfy their concerns that data might be lost within the state and federal agency domain. In addition, some agencies have exchanged documents over the years depending on who within a particular agency at a given time developed an interest in the vegetation monitoring efforts on Roan (e.g., datasheets originally housed at the U.S. Forest Service are now housed in the U.S. Fish and Wildlife Service office).

b. Bob Harvey's Briar Bashing Legacy Database

The majority of spatial documents (maps) were obtained from SAHC-Kingsport office. A few additional documents were obtained from the data originator (Bob Harvey) either during a meeting or through regular mail. However, the 1995 map was missing from Bob Harvey's records. After a considerable amount of searching through SAHC-Kingsport office files, this document was retrieved.

Associated data documents (e.g., cost-analysis records and descriptions of the mow areas) were located in relatively unorganized SAHC-Kingsport files and had to be processed further in order to exact relevant information. Due to the incomplete data documentation for some years, it is not possible to calculate a complete cost-analysis.

c. Black Berry Eradication Study Legacy Database

These data documents were located in the SAHC-Kingsport office.

d. Tom Kenney's 1999 GIS Legacy Database

Initially, an incomplete hardcopy (included only text) was obtained from the SAHC-Kingsport office. A complete hardcopy (included both text and electronic spatial files) was eventually obtained from Jamey Donaldson. An electronic copy of the spatial files was located in the SAHC-Asheville office.

3. Level of Concise Documentation and Metadata Associated with Them

a. Transect and Plot Dataset

In general, it was a challenge to locate and compile relevant information for metadata documentation because of the inherent nature of the dataset. Information from any individual source was incomplete and typically multiple sources needed to be consulted, and then revisited as new information became available. The primary characteristics of the dataset that made it difficult to work with was that the dataset was generated 10-20 years ago, spans 10 years and was not systematically documented with a standard methodology.

Waning Recollection
Due to the time that has elapsed since the data were collected, originators of the data often could not remember details. Occasionally information generated from interviews with originators of the data contradicted information presented in data documents or by other data originators. Originators of the data, who have remained more closely involved relative to others over the years with Roan projects, had a better recollection of details and therefore were able to supply more information. Greater than half of the data originators were forthcoming with information while others were more hesitant. Hesitation was probably a combination of varying degrees of involvement in the original projects, recollection, time constraints, depth of continued interest in original project goals, and basic personality. Meetings with the data originators produced the greatest wealth of information because dialog was more open and spontaneous with immediate follow-up, and documents being discussed were tangible on both ends of the correspondence.

Job Turnover – data acquisition, correspondence, and response time
Nearly all of the data originators are employed at other jobs at this time, and in some instances employment is several states away making acquisition of data documents and correspondence difficult. Along with these difficulties come time constraints on the part of the data originator. Slow response time to requests for information on project details was typical, and as a result requests for information from interviews were not always able to be presented in the most efficient and logical manner due to varying response times of individual data originators.

In addition, there was warranted concern surrounding the transport of original materials by mail in terms of the possibility of losing irreplaceable documents and therefore special arrangements had to be made to retrieve and return documents. Although many data documents have been retrieved and are now being photocopied and archived, some data documents still reside in the more distant geographic areas as the result of the geographic distance and time constraints imposed upon the part of individual data originators.

Extensive Processing – data documentation (organization and systematic documentation)
Once data documents were obtained, often extensive processing was needed in order to exact relevant information.

- Relevant documents occasionally had to be retrieved from unorganized files. That is, data files had not been well maintained over the years.
- Field methodology and/or data analysis procedures were sometimes not well documented. For example, most documents are unpublished and in the form of internal agency documents, and therefore would not have been peer reviewed, nor were they subjected to a standard format for systematic documentation of scientific data. Likewise, a notable number of documents were in draft verses final form and the most up-to-date draft had to be identified. Ultimately this means that some documents were incomplete on some level and/or had not been fully reviewed. As a result, data originators had to be contacted individually to fill in gaps in information. The interview process was very time consuming; interviews not only required the time on the part of the interviewer but also the data originator. However,

without the interview process, compiling the metadata documentation for this complex dataset would be considerably less complete than it presently is due to the time that has elapsed and the condition and location of the documents. On occasion it was difficult to compile relevant information from documents. That is, some or all information in the document was not accurate in the sense that some documents contained information on projected management plans instead of plans that were implemented (meeting minutes, projected study design, etc.). Also, decisions had to be made when documents were encountered that contained an abundance of information on rare plant occurrences. These documents had to be pulled and sent to respective State Natural Heritage repositories.
- Spatial files, although few existed, had to be further developed in order to extract relevant GPS location data.

Discrepancies and Inefficiency
As a result of insufficient recollection, poor data documentation, and slow response times with data acquisition (interview arrangements and/or data documents), discrepancies in information were generated and often data documents had to be revisited and data originators had to be contacted more than once. Although most discrepancies were eventually resolved some still remain and have not been answered to date. Where discrepancies existed, generalizations about the utility of the data had to be made.

Metadata documentation did not come without a considerable expenditure of time and effort. This is a very inefficient process to uncover the details and often the details cannot be determined years later. Data needs to be documented systematically the first time.

b. Bob Harvey's Briar Bashing Legacy Database
The processing time was extensive. An interview by phone and eventually a meeting were necessary to fully understand the spatial data files and associated data documents. Spatial documents (e.g., maps) had to be further developed (scanned, georeferenced, and then digitized), using GIS software because this information previously existed only in hardcopy form. A considerable amount of time was spent organizing the associated data documents (time sheets, expenditures, letters containing descriptions of mow areas, etc.). Associated data documents represent an incomplete set to date, and therefore this will hinder the compilation of thorough cost-benefit analysis.

c. Black Berry Eradication Study Legacy Database
Contact information for only one out of three data originators could be located, and thus information, extraneous from the data documents, is based solely on the one data originator that was interviewed. This data originator had a full understanding of the 1987 data collection efforts, and only a general understanding of the 1986 data collection efforts and the proposed study design. Due to time constraints imposed upon the one data originator interviewed, replies to information requests were slow in coming.

d. Tom Kenney's 1999 GIS Legacy Database
Information directly associated with the Roan grassy balds community had to be extracted from the report and spatial files. Because projection information associated with the spatial files was not provided in the hardcopy report, ArcCatalog was required to extract this information.

4. Condition of the Data Documents and Associated Materials

a. Transects and Plots Dataset
Overall condition
Most data documents, whether obtained from agencies and organizations or home residence, were in good condition with the exception of a few materials. One folder of documents had been exposed to mold in the basement of a private residence; however, these documents had not depreciated to a non-legible state. Also, a small portion of the transect and plot photographs collected in 1987 and 1988 had been damaged (see "Condition of datasheets and associated photographs" section).

Hardcopy Datasheets
Three sets of datasheets were located; however, only two of the three sets have been made accessible by this date. Paul Somers likely has a set that contains the original 1987 and 1988 datasheets. It is unknown at this time how complete these sets are because complete inventories have not been conducted. A complete inventory will require a notable amount of time due to lack of organization of the data sets and the amount of material that needs to be sorted. In some cases, groups of data sheets are clipped to a summary data sheet that contains information relevant for interpreting the underlying datasheets. A preliminary analysis shows that data sheets are mostly legible. Below is a very brief summary of the data sheets that have been located to date:

<u>U.S. Forest Service</u>
1987: Round Bald, Little Hump Mountain, and Bradley Gap
1988: Grassy Ridge, Big Yellow Mountain, and Hump Mountain
1993: Grassy Ridge

<u>Alan Smith</u>
1987: Round Bald, Jane Bald, Little Hump Mountain, Bradley Gap, and Hump Mountain
1988: Grassy Ridge Bald, Big Yellow Mountain, Little Hump Mountain, Bradley Gap, and Hump Mountain
1992: Round Bald and Jane Bald
1993: Big Yellow Mountain

Electronic Datasheets
At least one electronic set of the raw data spanning 1987, 1988, 1992, 1993, and 1994 was obtained from the U.S. Forest Service, and one electronic set containing the 1987 and 1988 raw data was obtained from Paul Somers. Any conversion of files from older file formats was undertaken by prior investigators (i.e., prior to metadata compilation in 2005-2006).

The electronic set containing data spanning from 1987-1994 has not been fully inventoried and will require a considerable amount of time to crosswalk with hardcopy datasheets once hardcopy datasheets are sorted. This data was entered electronically in 1994 from raw data sheets by U.S. Forest Service personnel (Chris Ulrey and Joe Noto). Physical parameters were not entered electronically in 1994, because there was only an interest in analyzing the vegetation parameters at the time. This electronic set is comprised of four Excel files: ALLROAN_plots.xls, ALLROAN_plots2.xls, RNDsum.xls, and ROANSPEC.xls.

There is uncertainty as to which of two files contains the most current version of the raw data for all years and areas. More specifically, ALLROAN_plots.xls is similar to ALLROAN_plots2.xls, however the data originator noted that it is likely that ALLROAN_plots2.xls (which contains more data than ALLROAN_plots.xls) is probably more current than ALLROAN_plots.xls. Below is a brief summary of the data contained in ALLROAN_plots2 file:

<u>USFS</u>
1987: Round Bald (all), Jane Bald (all), Little Hump Mountain (all), Bradley Gap (half), and Hump Mountain (half)
1988: Grassy Ridge Bald (all), Big Yellow Mountain (all) , Bradley Gap (half), and Hump Mountain (half)
1992: Round Bald (all) and Jane Bald (all)
1993: Round Bald (subset)
1994: Round Bald (subset)

In addition, the utility of RNDsum.xls is uncertain. This file appears to contain only Round Bald data that was queried from one of the larger, more comprehensive files (ALLROAN_plots.xls or ALLROAN_plots2.xls). However, this file should be crosswalked with ALLROAN_plots2.xls because it appears to have some serious problems. The data does not line up in some portions of the file, some data fields were not able to be defined, and it is possible that the data is representative of an output from an analysis instead of raw data. Data originators could not recall exactly what the data represented due to the time that had elapsed since developing the files.

ROANSPEC.xls contains information that defines the species and life-form codes recorded in the afore-mentioned Excel files: ALLROAN_plots.xls, ALLROAN_plots2.xls, and RNDsum.xls. This file appears to be intact.

The electronic set containing 1987 and 1988 raw data that was entered electronically in 1989 has not been inventoried to date and is contained in approximately five Dbase files. These files contain data on both vegetation and physical parameters. No distinction was made within the files to indicate which year (1987 or 1988) the data were collected on each bald. Below is a brief summary of the data contained in these files:

Paul Somers:

1987 and 1988: Round Bald, Jane Bald, Grassy Ridge Bald, Big Yellow Mountain, Little Hump Mountain, Bradley Gap, and Hump Mountain

Photographs

Although a nearly complete set of photographs of the transects and plots were taken, a 1996 inventory revealed that by 1996 some photographs were missing, some negatives were damaged from being stuck together or had emulsion damage, some prints were too obscure as a result of double exposure or were unexposed, some prints were duplicated, and some prints were without labels. As result of the 1996 inventory, existing photos are organized chronologically and by bald in three-ringed folders, and there are many photographs and negatives. However, the completeness of the photo set as noted in the 1996 data document has not been cross-walked to date with the hardcopy photo set to determine completeness due to time constraints. Below is a preliminary inventory of the photos:

Five folders with prints from the following bald areas:

1. Round Bald and Jane Bald
2. Grassy Ridge Bald
3. Big Yellow Mountain
4. Little Hump Mountain, Bradley Gap, and Hump Mountain
5. Hump Mountain

One folder with negatives from the following bald areas:

Round Bald, Jane Bald, Grassy Ridge Bald, Big Yellow Mountain, Little Hump Mountain, Bradley Gap, and Hump Mountain

Two small boxes:

1. duplicate prints from various bald areas
2. a couple of unlabeled prints

Species codes

Only a general portrayal exists for how species codes were derived through the years as a result of insufficient documentation and recollection as to what species codes were used 10-20 years after the individual studies were conducted, and the overall lack of using standard methodology to document species code in some cases, in combination with changes in species codes through the years,

Different species codes were used at different times, either unofficial species codes or codes based on different state or federal standards depending on the year that the data was collected and who collected the data. This is a classic example of why standard methodology is important, especially when dealing with a dataset that spans multiple years. As a result, lack of standard methodology through the years has resulted in incompatibility among datasets. Data originators who were interested in comparing data from multiple years and/or just entering data from original datasheets had to spend a considerable amount of time in the past cross-walking original datasheets with species list where they existed, and consulting the data originators. As data originators from earlier years have noted during this process, the lack of a standard methodology and poor data documentation of the species codes in general have resulted in a very inefficient problem hindering the efficient acquisition of the data, in a comparable format, for comparison.

The latest attempt to compare data and thus standardize species codes among years was conducted in 1994. Data from multiple years (1987, 1988, 1992, 1993 and 1994) was entered electronically, and in the process of entering the data electronically species codes were standardized among years. If the 1994 electronic version of the data is determined to be incomplete, then a considerable amount of time will need to be spent searching through original documents, notably the raw data sheets and species code lists where they exist in order to decipher data species codes from the different years.

b. Bob Harvey's Briar Bashing Legacy Database
Hardcopies of the maps exist for most years, except for 1992 when the data originator indicated that data was not collected because of his absence from the field site.

c. Black Berry Eradication Study Legacy Database
One set of hardcopy data sheets was obtained and through a preliminary inventory it appears to be complete and readily legible. No electronic files of the raw data exist.

d. Tom Kenney's 1999 GIS Legacy Database
One complete set of the spatial files electronic form, one hardcopy of the associated report, and an incomplete set of hardcopy spatial files was obtained. A preliminary review of the electronic version of the spatial files and the associated report appear to be complete and readily legible.

5. Access to the Datasets

a. Transect and Plot Dataset
Requests for data documents were presented first to agencies whom were identified as the primary repositories of Roan documents and secondly to individual data originators and their affiliated agencies or organizations. Data were obtained in the following manner: received as attachments through electronic mail; retrieved from personal meetings with the data originators, or affiliated agencies and nonprofit organizations; or mailed by regular mail.

It has been difficult to determine whether some agencies housed relevant documents because of discrepancies in information obtained through interviews with data originators associated with the same agency. In some cases, only through multiple requests with various persons working in affiliated agencies, was at least one person identified that knew about the relevant data documents housed within the agency.

Data documents came from several sources. Most data documents were obtained from the SAHC-Kingsport office, USFWS-Asheville office and Alan Smith (a primary data originator). Both, the SAHC and USFWS, are among the primary organizational entities that have continued to closely follow the monitoring efforts on Roan for many years. The USFWS had accumulated several Roan documents over the years. In addition, the USFWS recently obtained several data documents that had previously been housed at a USFS office due to renewed interests in curtailing the potential loss of documents. SAHC has accumulated many Roan data documents over several years because of the continued interests in the preservation of Roan, and thus was able to supply an abundance of relevant documents. Alan Smith has remained in close geographic proximity to the Roan area and has continued to closely following the monitoring efforts of the grassy balds over the years as an employee and volunteer.

Often special arrangements were made to meet with data originators to retrieve documents because of concerns that irreplaceable data documents may be lost through the regular mail system and to curtail costs associated with mailing documents. Meetings required travel on the part of the data originator and more often on the part of the data compiler. Once copies of the documents were made, originals were transported by the data compiler or others to the owners, or returned by regular mail. Mailing documents, especially boxes of material, can be an expensive process but on occasion this was the only choice in instances where the data originator lived many states away or several hours away.

Although many primary documents which are relevant to this legacy database have been obtained, not all relevant documents have been obtained to date. It has not been possible to obtain all data documents up to this time and there are indications that not all relevant documents will be uncovered because they have been misplaced over the years, no longer exist, and/or were not generated for some years (time sheets, areas mowed, etc.). For example, not enough data documents have been recovered to date to compile a full cost-analysis of the transect and plot project. Some documents are in draft version, and final versions have not been located and/or likely were never generated. Furthermore, personnel from the USFS, TNC, and Roan Stewardship Committee in addition to individual data originators have indicated that relevant data documents are presently housed in their work or home facilities but due to time constraints with compiling the documents they have not been able to provide documents to date. Overall, data documentation, to the extent practical and feasible at this point in time (10-20 years later), is fairly complete for most studies conducted during 1987-1996 in the sense that most data documents have been located and retrieved.

b. Bob Harvey's Briar Bashing

Most documents were obtained from SAHC-Kingsport office and Bob Harvey, and some documents were obtained from the Appalachian Trail Conservancy. Travel arrangements were made to retrieve data documents from SAHC files. After reviewing documents in the SAHC files, it was necessary to arrange a meeting with Bob Harvey for further clarification. The ATC was able to compile relevant documents and mail them by regular mail. All known spatial documents have been obtained but some non-spatial documents (time sheets and associated expenditures) were not able to be retrieved because they likely lost and/or for some years may not have existed.

c. Black Berry Eradication Study Legacy Database

Data documents came from the SAHC-Kingsport office. Travel arrangements were made to acquire data documents from both these sources. All known data documents have been obtained.

d. Tom Kenney's 1999 GIS Legacy Database

Data documents came from the SAHC-Asheville office and the USFWS-Asheville office. Travel arrangements were made to retrieve data documents from both of these sources. All known data documents have been obtained.

6. Degree of Digitization of the Datasets

a. Transect and Plot Dataset

The degree of digitization was minimal due to the (1) small amount of spatial data collected during the transect and plot project, and (2) problems associated with recovering spatial data by 2005 (nine years after it was collected).

(1) Only a few attempts were made to acquire spatial data using a GPS unit because availability of units were limited at the time the study was conducted, and when a unit was available it appears that data originators had limited knowledge on how to operate the unit and therefore were not able to collect spatial information or did not attempt to collect spatial information.

A less well-known attempt was made to GPS transects and plots located within the confines of the goat grazing paddocks in 1992-1994. However, again, due to the lack of knowledge on how to operate a GPS unit and the low accuracy associated with the GPS in the mid-1990's, spatial data was not recorded even though a GPS unit was available. Again, it appears that a systematic field methodology was not in place for documenting spatial data.

There is also concern as to why alternative techniques were not used, if GPS was not available nor understood, to systematically document the spatial context of the work done (e.g., possibly surveying).

(2) An attempt was made in 1996 to GPS all transects and plots across the balds complex during a relocation effort. However, the GPS data acquired during the 1996 field season has been determined to be of little use beyond a general portrayal of what is to be assumed were point locations of plots. The following problems were encountered:

- Some spatial files containing GPS data were lost during the field efforts and/or were not located in electronic files although data originators had made an attempt to place the data in designated electronic files and folders. Data originators were uncertain why files were not available on data disks or specifically why they were lost during field efforts, other than the fact that the GPS unit was forgotten on at least one field day.
- Accuracy levels of the GPS files could not be determined because data documents did not describe the 'standard collection methods' used. It appears that no systematic field methodology had been developed nor provided by the USFS to document spatial data.
- During the time that GPS data would have been collected (approximately the mid-1990's), GPS Selective Availability was still turned on and considerably fewer satellites would have been present relative to the number that are present at the current time.

Although an attempt was made to differentially correct points in 2005 from base station files at the North Carolina Geodetic Survey (NCGS), not all base station files were available nine years later. For files that data was available, data was differentially corrected using Trimble Pathfinder Office software, and for each respective GPS file the points were grouped. Where NCGS base station files weren't available for specific GPS files, the differential correction step was skipped. The files, assumed as intended to be point data locations, were then averaged or grouped so as to obtain a more concise point location. ESRI ArcGIS was then used to re-project the files into the U.S. Forest Service standard Geographic Datum and Projection of NAD83 UTM 17N. As a result, the spatial files should be used for general portrayal and general information only.

b. Bob Harvey's Briar Bashing Legacy Database
Hardcopies of maps were digitized (1989 – 2004, with the exception of 1992) in 2006 to determine the utility of the data.

The basemap for the maps was adapted by Bob Harvey from a USFS outline map of the area of Round Bald and Engine Gap. No specific information is known about the USFS basemap that was provided to Bob Harvey regarding the accuracy, time period it was collected, or methodology used. The original electronic file, after speaking with USFS personnel, has been determined to be lost. It is assumed that it is was a Trimble GPS unit file as on the original copy of the USFS basemap given to Bob, an electronic computer file path showing the apparent location of the file on the 'C' drive of a computer, in a folder labeled PFPRO, with a .COR file extension, which likely indicates it was generated using Trimble equipment and software due to the file naming conventions.

Managed areas were approximated and were drawn and shaded in to show the areas of briar cutting done each season by SAHC volunteers, contract persons, and USFS. On maps with GPS latitude/longitude coordinates for cut-over areas, a GARMIN Model GPS III Plus unit was used (2001-forward). The method used to show the latter on the map consists of extremely general hand-drawn approximations for GPS coordinates taken and managed vegetation areas, and were not produced with any quantifiably accurate method. GPS data coordinates were read at the location, and then recorded in hardcopy format.

Per notes on copies of Bob's maps, the landmark positions and descriptions were obtained in 1994 and 1999 by rather crude surveying—one person working alone with plastic surveyors chain and Brunton compass. It was noted on the maps that it is hoped that the error of positioning is generally not greater than 50 feet for these.

c. Black Berry Eradication Study Legacy Database
No spatial data was collected.

d. Tom Kenney's 1999 GIS Legacy Database
A total of seven spatial files were exacted from this dataset. These files did not need to be further developed (e.g., digitized). However, what is not known is the percentage error associated with vegetation areas delineated by the author.

7. Clarity of Defining What the Data Represents – e.g., How "Percentage of Plot Cover" was Determined

a. Transect and Plot Dataset

During the period from 1986-1996, the data originators of the transect and plot data used ocular estimates to obtain percent cover within meter-square plots. Data originators (comprised of a team of botanists) determined that use of ocular estimates to obtain percent cover was the best method to use for the purposes at hand. There was a need to maximize the area covered across the balds and this method allowed enough time to cover all balds within the time allotted, although it was still an intense effort. Data originators were aware of the other methods (line intercept method, etc.) and the repeatability and accuracy associated with this methodology; however, due to the sampling intensity associated with other methods, sampling at the proposed scale with other field methods would not have been feasible.

Multiple field collectors would have collected percent cover estimates over the years. Notably, two different sets of field collectors would have sampled over the years: two primary field collectors would have worked closely together collecting the 1987, 1988, 1992 and 1993 data across multiple grassy balds, and 3-4 primary field collectors would have worked together collecting the 1992, 1993, 1994 and 1995 data within the confines of the goat paddocks. The data originators that sampled during 1987-1993 across all balds noted that they made a concerted effort to work closely together on arriving at the similar estimates. Field collectors would come up with a cover number and negotiate and change their estimate or average their numbers. However, it is unclear at this point in time how closely the data originators, who sampled during 1992-1995 within the confines of the goat paddocks, worked at generating similar estimates.

Data originators have noted that the data is not obsolete due to the use of ocular estimates, wherein measurements of percent cover were recorded as absolute measurements and not as a cover class. Some researchers would argue that use of ocular estimates only documents presence or absence. However, if there is sufficient concern over the potential error associated with ocular estimates of percent cover between different observers or among years then estimates can be converted to cover classes when comparing data.

8. Level of Communications Among Project Partners, Biologists, Resource Managers and Scientists

Information exchange among all of the partners involved with stewardship and management of the Roan massif tends to be informal and on an as-needed basis, with the possible exception of communications taking place in conjunction with a consultation pursuant to the Endangered Species Act (between the USFS and USFWS), a Forest Plan revision, or impact assessments associated with NEPA compliance. There is no single trigger that prompts the exchange of information; every information exchange is likely prompted by a different need. Virtually everyone involved in this collaborative has been involved, to some degree, with the exchange of information.

a. Transect and Plot Dataset

Initiation

This project was initiated by Dr. Paul Somers, the State Botanist with the Tennessee Natural Heritage Program, in 1986. Dr. Somers provided the cost-share grant that would allow for the collection of baseline vegetation data through installation of permanent transects and plots (as well as associated collection of baseline data on rare plant distributions and small mammal populations) across the Roan grassy balds. Personnel from a total of eight different agencies and organizations worked together on the designing the study, writing the grant proposal, providing financial support, and collecting the 1987 and 1988 field data. Cooperators included the following:

- U.S. Forest Service, Cherokee National Forest, Unaka District
- U.S. Forest Service, Pisgah National Forest, Toecane District
- Tennessee Department of Conservation, Ecological Services Division

- North Carolina Department of Agriculture, Plant Industry Division, Plant Conservation Program
- North Carolina Department of Natural Resources and Community Development
- Division of State Parks, Natural Heritage Program
- U.S. Fish and Wildlife Service, Endangered Species Field Station, Asheville
- Southern Appalachian Highlands Conservancy
- Wake Forest University, Department of Zoology

Continuation of efforts throughout the years
From 1989-1996, the impetus for implementing specific projects came from collaboration of partners and others with management responsibilities and interests in the natural communities of Roan.

More specifically, the Roan Scientific Advisory Group developed out of the initial efforts of the multi-agency funded project designed to collect baseline data across the Roan grassy balds. This group, comprised of biologists and scientists, was responsible for the balds management planning, implementation and monitoring efforts that occurred on Roan. The advisory group identified and established the primary values on a bald by bald basis, and recommended management tools, base on this information, documented in the balds management plan (See USDA Forest Service 1991 for details[5]).

SAHC's Roan Stewardship Committee, comprised of the same agency, institutional and organizational partners, has served to carry forth these efforts in recent years. Throughout the years, SAHC has continued to play a pivotal role to ensure that efforts are in place to facilitate management by assisting in the provision of management funds, executing management activities, conducting and coordinating field assessments, organizing conferences, and arranging ad hoc meetings.

Sharing of documents
As previously noted, various documents have been generated as transect and plot data was gathered and analyzed over the years by multiple partners. The primary avenue for sharing these documents was among partners that participated in the Roan Stewardship Committee meetings. However, without the existence of a central repository for data documents partners did not have a systematic method for acquiring, storing, and assessing the multitude of documents that had been produced. Likewise, some documents might have been slatted as internal documents so these types of documents would not have been made available to all partners, unless they were specifically requested.

It appears that the same raw data would have been entered on different occasions as a result of lack of communication between partners.

<u>1989</u>
Around 1989, Paul Hamel entered the 1987 and 1988 plot data, line intercept data, data on physical parameters for all balds electronically.
Around 1989, Michael Schafale entered the 1987 and 1988 plot data for all or a portion of the balds electronically. Michael Schafale noted that he gave his electronic files to Steve Simon.
Around 1989, Steve Simon entered at least a subset of the 1987 line intercept data for Round Bald electronically.
<u>1994</u>
In 1994, Chris Ulrey and Joe Noto entered the 1987, 1988, 1992, 1993, and 1994 plot data and line intercept data for all balds electronically.

Similarly, analysis was duplicated in one instance because of lack of communication between partners. A community classification analysis was run using the 1987 and 1988 data by Paul Somers and Paul Hamel in 1992 and also by Chris Ulrey and Steve Simon in 1994.

---

[5] U.S. Department of Agriculture Forest Service. Roan Mountain Highlands Vegetation Management of the Grassy Balds: Environmental Assessment. 1991.

b. Bob Harvey's Briar Bashing Legacy Database

On the ground

The various management partners have continued to work together to identify the annual mowing needs. Management activities, such as mowing, have come through the USFS, SAHC and ATC. SAHC has maintained a consistent level of mowing effort through the years and has taken on a role of coordinating the mowing with multiple other agencies that conduct studies, monitor populations, or implement on the ground management on the Roan grassy balds. On an annual basis, SAHC has worked closely with various partners and their crews to coordinate mowing efforts - USFS crews, various volunteer-based SAHC crews, the North Carolina Department of Corrections Bridge Crew, and ATC-sponsored crews.

Sharing of documents

Bob Harvey (a SAHC volunteer) has compiled the most extensive records on the mowing efforts on Round Bald and in Engine Gap over the years at the request of the SAHC Stewardship Director, Judy Murray, who identified the need to document the annual mowing efforts. Bob has provided maps to the various partners on a nearly annual basis during Roan Stewardship Committee meetings.

Overall

Although various partners have participated in annual mowing efforts on the ground and discussed the annual mowing needs on occasion during ad hoc meetings on the balds or during formal meetings (e.g., Roan Stewardship Committee meetings), there is still a lack of partner communication and participation in terms of identifying a systematic way of monitoring the woody encroachment on the grassy balds communities.

c. Black Berry Eradication Study Legacy Database

The impetus for the study came from discussions, among project partners, surrounding the need to evaluate different management tools. USFS personnel took the lead on the project in terms of final design, funding and field implementation and assessment.

d. Tom Kenney's 1999 GIS Legacy Database

This compilation developed as a result of collaboration between the various partners previously mentioned. This project was financial funded by The Nature Conservancy and the Cannon Foundation. Tom worked closely with the various partners to identify, update and compile relevant spatial and biological information on natural faunal and floral communities from data sources (e.g., files, databases) within the cooperating agencies and organization in an effort to generate unified GIS map layers that could be used by the various partners for conservation purposes. However, there was limited feedback from partners during the development of this dataset. When Tom requested that the various partners review this data prior to submitting a final report, only one partner submitted edits.

9. Summary of How the Data and Resulting Information Has Been Utilized

a. Transect and Plot Dataset

Past Use

Data collected over the years has been used in various capacities at different times and on different levels:

- The 1987 and 1988 data has been analyzed to document the baseline conditions of all grassy balds within the Roan grassy balds on two occasions (in 1992 and 1994) in order to evaluate the effects of future management (or natural processes) on grassy balds vegetation.
- A portion of the data (small subsets of data from Round Bald and Jane Bald) has been used to evaluate the effects of different management techniques (mowing and goat grazing) on the grassy balds vegetation.
- Some data collected over the years were never analyzed. [This may be the result of not clearly identifying objectives prior to data collection.]

- Photos taken in 1987 and 1988 were invaluable in relocating transects and plots during the 1990 and 1996 relocation efforts.
- Field notes from the 1990 relocation effort were invaluable in relocating transect and plots during the 1996 relocation effort.
- Data collected during the 1990 and 1996 relocation efforts provided a basis for determining whether it was feasible to revisit and resample vegetation at transects and plots in subsequent years. Re-staking intermittently over the years had the potential to be invaluable in of maintaining the permanency of the transect and plots. That is, a greater number of plots and thus transects could be re-sampled over the years than would have otherwise been if re-staking did not occur.

Compatibility and Comparability

The level of compatibility with existing data and potentially future data is presented below:

- Because of the variability in percent ocular estimates among observers across years, comparability of data will need to be considered on a case by case basis in relation to the specific project objectives and analysis.
- The scale on which this study design was implemented (sampling at approximately 899 plots) has not been feasible on the same scale since the initial installation because of financial and manpower limitations, and will not be likely at a future date. This is further complicated with the inability to precisely relocate a majority of the original transects and plots.
- Transects and plots have been difficult to locate due to oversights in study design and lack of standard field techniques, as well as Quality Assurance so as to ensure consistency during development of the initial study design. The longer transects and thus associated plots were difficult to relocate because the inherent difficulty in maintaining the original bearing for the entire length of individual transects. In addition, the relocation of transects was difficult because declinations on compasses used were not standardized among all field collectors.
- The exact location of a substantially large percentage of transects and plots are not known because they were not georeferenced over the years and, as a result, comparison with actively managed areas (or natural succession) is limited when attempting to compare data.
- The exact location of goat paddocks over the years is not known because paddocks were not georeferenced, and therefore the utility of the data is limited when attempting to compare data.
- The exact location of areas mowed over the years is not known because these areas were not georeferenced, and a result the utility of the data is limited when attempting to compare data.

b. Bob Harvey's Briar Bashing Legacy Database

Past Use

The use of the data has been minimal to nonexistent. Bob likely consulted it for his personal use prior to annual mowing. However, the Southern Appalachian Highlands Conservancy personnel, who have taken the lead on organizing the annual mowing efforts, relied on their own field assessment of the woody coverage over the years. SAHC would have indirectly consulted associated data documents which would have provided information on logistical planning (e.g., equipment needs and expenses) for the mowing season the following year. It appears that the management partners that participated in mowing efforts, USFS and ATC, did not make use of the spatial data either. Specifically, it appears that data analysts within the USFS, who analyzed data collected from transects and plots within the confines of mowed areas, did not incorporate mowing records (frequency and intensity) into their analysis even though mowing records existed for some areas. It is uncertain why data analysts did not incorporate this important data into their analyses.

Compatibility and Comparability

The level of compatibility with existing datasets is limited in that only a general interpretation of the data would be possible. In terms of comparing and integrating existing transect and plot data, data analyst would need to crosswalk the subset of data proposed for analysis with mapped coverage of mowed areas. However, only a general interpretation would be possible for several reasons.

- Mowing has not been uniform due to patchiness of woody vegetation and on the ground mowing intensity. Although mowing is uniform in some areas, more so in areas where the track mower is used, it is not uniform in all areas, but instead often the denser blackberry patches are targeted and mowed by the field crews on a case by case basis from year to year. Although mowing intensity (in terms of mowing below the lowest leaf on blackberries) is stressed prior to mowing efforts, there can be considerable variation between field personnel.
- GPS coordinates are not available for most, if not all, transects and plots where data were collected and therefore transect and plot locations relative to the mowed area mapped would have to be estimated.
- Accuracy and completeness of the map coverages is uncertain. It has been determined through digitization of the map coverages that boundaries of mowed areas are off by approximately 50 meters for some years. In other words, mowing intensity and frequency would have to be estimated due to the lack of geo-referenced transect and plot locations, lack of uniformity in mowing, and relatively large amount of error associated with boundaries of map coverages.
- The costs associated with this study design have not been compiled to date, and thus the feasibility of conducting a comparable study in terms of financial limitations cannot be evaluated at this time.

c. Black Berry Eradication Study Legacy Database

Past Use

It is uncertain to what extent this data has been used. It is likely that the use of the data documents generated from this study has been minimal to nonexistent. In addition, it should be noted that herbicide application on the western balds (Round, Jane & Grassy) was considered unacceptable by USFWS personnel (Nora Murdock) and other biologists during the 1980's and 1990's.

Compatibility and Comparability

- Percent cover would have been collected using ocular estimates. This sampling technique should be taken into consideration when comparing data across years because different field collectors would have been present between 1987 and 1988.
- The treatment area was not geo-referenced so the likelihood of relocating the exact plot again to determine the effects of the long-term responses and/or conducting a similar study in the same area at a future date is not possible.
- Because of the lack of uniformity in herbicide application (e.g., spillover into adjacent treatment areas), and the time of year that the herbicides were applied, only a general interpretation of the effects of herbicide application can be inferred.
- The scale on which this study design was implemented (thirty-six 20 X 20 meter plots) would be feasible at a later date. However, it was noted by data originators that because of the relatively large plot sizes, the effects of herbicide application and mowing on individual grass and sedge species were not differentiated.

d. Tom Kenney's 1999 GIS Legacy Database

Past Use

It is likely that the use of the data documents generated from this study has been minimal to nonexistent. It appears that only a few of the partners (The Nature Conservancy; Jamey Donaldson, contract biologist; and SAHC) are aware of the existence of this legacy database and thus its potential in addressing adaptive management issues, although the various partners were asked to review a draft version of the data during a 1999 Roan Stewardship Committee Meeting, and the final report has been provided to the various partners since. SAHC personnel from the SAHC-Asheville office have consulted the data as a land parcel locator map and for identifying conservation values on particular properties when developing potential land purchases and conservation easement projects.

Compatibility and Comparability

Specific examples of how the data is compatible and comparable, is unknown, at this time. However, it appears that the data contained therein may be comparable to data maintained by the various Roan partners. That is, many of the data compiled in this dataset was derived from existing data which was obtained from a variety of databases maintained by the partners (NCNHP, TNNHP, etc.). In terms of comparison with future analytical tools, ArcView was used to compile spatial data and this analytical tool (although later versions have been developed) is still in use at the present time. The data, however, can be readily integrated into more current versions of ArcGIS as well.

10. Bottom line: how did the datasets make a difference?

a. Transect and Plot Dataset

Plot-Transect 1987-1988 Installation

The 1987 and 1988 data served as a good (but impractical in the long-term) baseline upon which to measure future changes in community composition in response to management activities (and natural succession).

This dataset was the first attempt to systematically classify the vegetation communities on the Roan grassy balds and thus established baseline information for measuring change in the grassy balds community for which the effects of management (mowing and goat grazing) could be evaluated. In addition, documentation of the Roan grassy balds vegetation of the late 1980's, allows qualitative comparison to the vegetation types described by Brown from the late 1930's and Mark from the late 1950's.

Transects and plots were photographed when they were originally installed. These photographs provide additional, primarily qualitative, information as to the (vegetation) composition of the grassy balds prior to the initiation of vegetation management activities.

Because of the large spatial scale across which data were collected, data originators have suggested that it might be more realistic to revisit and resample a subset of the original plots and transects across the balds, if it is determined that plots and transects should and can be revisited at a later date. It might be more realistic to collect only line intercept data (verses plot data) because of financial constraints, as long as data collected is determined to be valuable and thus meet project objectives. This decision should be preceded by devising a robust scientific analysis of the monitoring objectives related to adaptive management and science. What are the research questions, such as "What are the drivers to sustain vulnerable populations at the community and the species level?" and "What are the cumulative effects from threats to sustainability such as invasion by woody plants, invasive species, pests, pathogens and climate change?"

Plot-Transect 1990 Relocation

This field effort documents data on the transects and plots that were located and restaked in 1990. This field effort provides an idea of how feasible it was to relocate plots and transects, and how many permanent markers were intact as of summer 1990 (two to following the original installation of the permanent markers).

Plot-Transect 1996 Relocation and Plot-Transect 1996 Relocation GIS

This is the only field effort conducted wherein plots and transects were geo-referenced. This field effort also provides information on the number of plots and transects that were able to be relocated and restaked as of 1996 (six years after the 1990 relocation effort, and seven and eight years post-installation of the permanent markers).

Accuracy levels of the GPS files cannot be determined as original notes did not describe the 'standard collection methods' used. In addition, when the files were taken, GPS Selective Availability was still turned on during 1996. The few files that contained GPS location information that could be differentially corrected in 2005 represented only a very general portrayal of transect and plot locations.

Plot-Transect 1987 and 1992 Mowing Analysis

The effects of hand-mowing on the vegetation composition of Round Bald can be inferred from the analyses of the 1987 and 1992 data. Interpretation of the data may be best presented as a general trend rather than specific quantitative measures because of limitations inherent in the sampling methodology, such as: estimates of percent (vegetation) cover; and lack of consistent and precise documentation of the location, timing, and nature of managed (mowed) areas.

These limitations have prompted several of those involved with the original data collection effort, as well as those who have conducted analyses of these data, to suggest that remotely sensed imagery may be a more appropriate tool for monitoring change in vegetation composition in the future.

Plot-Transect 1992-1993 Field Collections
In general, the 1992 and 1993 data was used in various analyses to monitor changes in vegetation on portions of Round Bald and Jane Bald in response to different management activities (mowing and grazing) by comparing data collected across various years. However, a significant portion of the data appears to have never been used.

Plot-Transect Community and Mowing Analysis
The vegetation composition classification analysis of the 1987 and 1988 data provides baseline (pre-management) conditions of the vegetation composition of the grassy balds complex of Roan for which the effects of management actions can be evaluated. In particular, the effects of hand-mowing on the vegetation composition of Round Bald and Jane Bald can be inferred from the analyses of the 1987 and 1992 data. Interpretation of the data may be best presented as a general trend rather than specific quantitative measures because of limitations inherent in the sampling and analytical methodologies, such as: estimates of percent vegetation cover; lack of consistent and precise documentation of the location, timing, and nature of managed (mowed) areas; inherent subjectivity of vegetation community classification analysis; and flaws in TWINSPAN software used.

Although the present sampling design is considered an adequate field design by many researchers, those limitations noted above have prompted several of those involved with the original data collection effort, as well as those who have conducted analyses of these data, to suggest that (1) cover classes be used in the future because this collection method allows for repeatability and consistency, (2) future analysis consider converting existing data to cover classes for more accurate comparison between years, (3) both physical and vegetation parameters be included in future analyses, (4) other types of field sampling (e.g., point intercept method) and statistical analyses (e.g., Cluster Analysis, ordination technique) be considered, but only after specific monitoring objectives are defined.

Plot-Transect Goat Grazing Study
In general, the effects of goat grazing on the grassy balds vegetation of Round Bald can be inferred from Steve Simon's analysis of the 1992-1995 data and associated documents cited herein. Interpretation of the data may be best presented as a general trend rather than specific quantitative measures because of limitations inherent in the sampling and analytical methodology, such as: estimates of percent (vegetation) cover and lack of consistent and precise documentation vegetation sampling dates and data origins among years that would allow for a full understanding of comparability of data between years.

Those limitations mentioned have prompted the concern that the effects of goat grazing on the grassy balds be explored further. However, it should be noted that from a preliminary interview with U.S. Forest Service personnel, this management technique may not be a priority due to the results of the U.S. Forest Service cost analysis study previously mentioned in Part 2 of this section. It was determined that the profit margin was too low for local permittees to pursue grazing activities due to the expenses and work associated with maintaining goat herds on the balds.

b. Bob Harvey's Briar Bashing Legacy Database
Bob Harvey's maps are the most thorough documentation of the mowing efforts implemented on the grassy balds during 1989-2005. In particular, the GPS spatial data that was collected from 2001-2005 is the only usable spatial data showing treatment areas. Though the accuracy is questionable, this is the most thorough record as stated and it serves as a starting-point model for more carefully planned future management efforts and documentation.

There is a need to identify a standard methodology for systematically documenting the boundaries of mowed areas, as well as intensity of the mows. And, thus data on mowed areas can readily be available for incorporation into data analysis. It is anticipated that a natural resource-grade GPS is needed to systematically document boundaries of mow areas. There is a need to ensure that adequate technical assisted is provided to field personnel where needed. In addition, cost-analysis records should be maintained in order to further evaluate the cost effectiveness of the various mowing techniques and monitoring tools. A dollar amount will illustrate the need for funding.

c. Black Berry Eradication Study Legacy Database
The overarching conclusions indicate that herbicides can be effective at controlling blackberry; however, this is to the detriment of killing desirable species. This should serve as a model for a starting place for further experimentation with herbicide application.

Interpretation of the data may provide general indications, rather than a quantitative measures, as to what the potential effects of herbicides (Garlon 4 and Round-up) and mechanical hand-mowing are because of the limitations inherent in sampling and analytical methodologies, such as: estimates of percent vegetation cover; short duration of the study; lack of full statistical analysis of the data; and lack of uniform application of herbicides across study plots. Conclusions drawn from the study should be a starting place for further herbicide experimentation, and herbicide application in combination with mowing (Alan Smith, personal communication-2005).

Given the limitations of the data and suggestions presented by one of the data originator, herbicide application should be explored further as potential tool for reducing woody vegetation on the grassy balds (Alan Smith, personal communication-2005).

d. Tom Kenney's 1999 GIS Legacy Database
A variety of existing and derived spatial data was compiled to provide a landscape tool for agency, organizations, and/or local groups interested in conservation planning on Roan.

Caution should proceed when using the data because percent errors associated with delineating boundaries is unknown. Despite potential limitations associated with boundaries, this dataset provides a landscape-level conservation design for Roan that can be consulted by Roan partners planning projects on many levels e.g., land acquisition, identification of important habitat types, rare species associated with various habitat types, preservation of faunal corridors, etc.

11. Quantifying the Time and Effort Spent (Time, Money, etc.) to Compile This Information; Metadata Documentation; and Design, Creation and Testing of the Toolkit

This section serves as an attempt to quantify the amount of time, effort, and most importantly, money that had to be spent to recover, document, and make available these identified priority legacy databases related to Roan. Besides the detailed descriptions of the problems associated with data management in the latter sections, quantifying the costs should serve to further stress the importance of proper planning and data management from the outset of an initiative.

The below numbers are based on approximate estimates (as close as possible), as well as time sheets, subcontract records, etc. We have confidence that they represent a fairly accurate portrayal of the resources spent (time, money, etc.) during this project year.

Monetary Expenses
A total of **$68,392.00** has been spent on this project. Funds supported costs (**$62,150.00**) associated with five project personnel (Tom Burley, Judy Murray, John Peine, Nora Schubert, Mary Thompson) and costs ($6,242.00) associated with supplies and materials (aerial photography processing and software, Toolkit expansion equipment (a Natural-Resource Grade GPS unit), general office supplies, travel, indirect administrative support) (Appendix B).

Hours Invested
A total of **3564.8** hours were spent by primary and secondary contributors compiling information for this Roan Data Management Project. The hours have been broken down by primary and secondary contributors (Appendix B).

2005-2006 Primary Contributors – Total Hours and Approximate Allocation of Time
Primary contributors consisted of personnel directly involved in data compilation and fulfillment of project objectives (e.g.,Tom Burley, Judy Murray, John Peine, Nora Schubert, Carolyn Wells, Mary Thompson). A total of **3375** hours were invested by primary contributors (Appendix B). Below is an overview of the tasks and approximate time associated with the compilation of information that was provided largely by the primary contributors. Approximate percent of time spent on each major task is presented, and tasks with asterisks (**) denote those tasks that required a significant amount of time relative to other tasks.

a. Project progress weekly (10%)
- Compiled weekly reports and expenditures to monitor project progress
- Management of general administrative overhead on project

b. Compiled a bibliography on the ecology of Roan and the history of adaptive management of the grassy balds (20%)
- **Searched various library sources
- **Developed database for managing bibliography

c. Created a catalogue and list serve of all key players and their roles (5%)
- Searched Internet sources and corresponded with data originators
- Developed database for managing catalogue of key players and their roles

d. Identified and located 'legacy databases' (20%)
- **Corresponded with data originators and agencies (and organizations) to acquire data documents and information
- **Arranged meetings with various data originators and agencies (and organizations) to acquire data documents and information
- **Searched and sorted agency and organization files to extract relevant data documents

e. Completed training in the following subjects (15%)
- USFS GPS guidelines
- Researched national protocols through correspondence with USFS and USFWS personnel and general literary searches
- **USGS NBII training in all aspects of the Data Management Toolkit
    - Attended professional Toolkit training sessions
    - Attended informal meetings with trained NBII-SAIN partners
    - Developed Toolkit and applied Toolkit to priority 'legacy databases' to the extent possible

f. Applied the Data Management Toolkit to the designated priority 'legacy databases' (30%)
- **Database documentation/Metadata compilation.
    - Developed a modified metadata form to facilitate metadata compilation
    - Extracted relevant information from data documents
    - Interviewed many of the data originators
    - Incorporated relevant information into metadata records
    - Attended a USGS-NBII sponsored metadata training workshop
    - Corresponded with USGS-NBII and project partners during final review and compilation of metadata records

- Digitization/data entry

- Digitized maps developed by Bob Harvey (UT and NBII-SAIN personnel)
- Corresponded with USFWS GIS specialist, SAHC-Asheville personnel, and ETSU GIS specialists regarding identification and digitization of imagery

- QA/QC analysis
  - Literature searches were performed to address accuracy of field techniques
  - Interviewed data originators and people knowledgeable of field and analytical methodologies used in regard to specific metadata files

- *Archiving both digitized and paper databases
  - Established a Roan archive agreement with ETSU-Archives of Appalachia
  - Corresponded with ETSU-Archives of Appalachia to address various aspects of project
  - Compiled and delivered documents for photocopying and archiving (initiated)
  - Requested permission from various partners to archive materials (initiated)

- *Geo-referenced data where appropriate/possible
  - Processed GPS files from Bob Harvey's records and the 1996 relocation using North Carolina Geodetic Survey base station files
  - Capabilities of a natural resource-grade GPS and associated software were explored during the 2006 transect and plot relocation effort

- Multi-agency adaptive management of the grassy balds
  - Researched and evaluated the roles of the various agencies and organizations by identifying priority legacy databases and through interviews (e.g, level of communication between partners, partner participation in Roan projects)

- Forest planning
  - Compiled summary of most current USFS plan
  - Correspondence with various people knowledgeable of the new USFS planning process

- Potential to create new clients to USGS NBII applications
  - Requests for participation was presented and sent to the Roan partners

- Evaluation of the Toolkit for application to data management and applied science initiatives
  - Final report compiled to evaluate the relevancy of the Toolkit to adaptive management by evaluating the effectiveness and utility of the legacy databases in the context of data management

2005-2006 Secondary Contributors

Secondary contributors consisted of the data originators and a number of other individuals who were able to provide supplemental information or support. A total of **189.8** hours were contributed as in-kind or paid services by these secondary contributors (Appendix B). The majority of the in-kind hours were supplied by those who participated in interviews related to specific metadata files, and secondarily related to correspondence associated with GIS technical support and data documents, data document compilation, one field visit to evaluate the feasibility of relocating transects and plots, US Forest Service planning, and archiving.

Appendix A - Bibliography Research 2005-2006

In 2005, an initiative was begun to compile a bibliography on the ecology of Roan and the history of adaptive management of the grassy balds to facilitate the archive process. Greater emphasis was placed on the history of adaptive management of the grassy balds due to time constraints. Despite time constraints, a fairly comprehensive Roan bibliography compilation is underway. The following is a summary of the number of references compiled to date:

- 885 references
- 114 references with "bald" in the title
- 156 references with "Roan" in the title
- 59 references for which SAHC has a hardcopy

The primary source of the literature references were compiled from records of the following agencies, organizations, and people:

- SAHC's comprehensive Roan bibliography that was compiled by Allan Trently in 2000
- boxes of data documents from the SAHC-Kingsport office
- the National Park's NatureBib database recommended by Kristin Johnson with the Great Smokey Mountains NP
- Tom Blevins at Mount Rogers NP
- Jerry Nagel, retired ETSU professor, who took the initiative in the past to begin archiving documents at the ETSU-Archives of Appalachia
- Jamey Donaldson's personal bibliographies

The secondary sources of literary information that were either searched and/or at least identified as potential repositories for Roan references are as follows:

1. World Wide Web
   a. Google Scholar
   b. Online Journals
      - Journal Access Links
      - NRS Online Journals
      - ESA Journals
      - SORA (Searchable Ornithological Research Archive; access is possible)

   c. Online Sites
      - Great Smokey Mtns NP
      - Virginia – Natural Areas Biblio
      - Peter White's webpage
      - USFS Publications Database
      - general search using key words
      - The Nature Conservancy
      - TWRA
      - TN Natural Heritage
      - NC Natural Heritage
      - NatureServe
      - NBII-SAIN
      - Mount Rogers, Jefferson NF
      - Friends of Roan Mtn

2. University Libraries – card catalogs, archive collections, journal articles, dissertations and thesis
   a. East Tennessee State University – Johnson City, TN
   b. Appalachian State University – Boone, NC
   [searches WNCLN (Western North Carolina Library Network) and thus combines ASU, WNC and UNC libraries]
   c. Wake Forest University – Winston-Salem, NC
   g. Duke University – Durham, NC
   h. Western Carolina University – Cullowhee, NC
   i. University of North Carolina – Asheville, NC
   j. North Carolina State University
   k. Mars Hill College – Mars Hill, NC

l. University of Tennessee – Knoxville, TN
m. University of Kentucky

3. State Libraries
    - TN State Library and Archives
    - NC State Library and Archives

The following key words were used during literary searches:

- *Primary - general*
    - Roan
    - Bald
    - Grassy bald
    - Mountain bald
    - Grassland
    - Appalachian
    - Southern Appalachian
- *Secondary - general*
    - Vegetation
    - Natural
    - Plant
    - Botany
    - Mt Rogers
    - Great Smoky Mountains
    - Vertebrate
    - Invertebrate
    - Zoology
    - Animal
    - Blue Ridge Parkway
    - High elevation community
    - Mitchell County
    - Carter County
    - Western North Carolina
    - Tennessee
- *Researchers*
    - Weigl, Peter**;** White, Peter; Peet, Robert; Knowles, Travis; Thompson, Cecil; Smith, Alan; Nodvin (early days); Gersmehl (early days), etc.

Appendix B - Cost Analysis for the Example Legacy Databases

The following table shows the contributors and their associated time and costs associated with compiling information for the Roan Mountain Data Management Project during 2005-2006, as indicated in section 11.

	Person or Item	Affiliation	Tasks	Hours Contributed	Costs
Primary Contributors					
	Tom Burley	U.S. Geological Survey-Southern Appalachian Information Node - Knoxville, TN	various (see section 11)	1200.00	22000.00

	John Peine	U.S. Geological Survey-Southern Appalachian Information Node - Knoxville, TN	various (see section 11)	510.00	20000.00
	Carolyn Wells	U.S. Fish and Wildlife Service - Asheville, NC	various (see section 11)	300.00	In-kind services
	Judy Murray	Southern Appalachian Highlands Conservancy - Kingsport, TN	various (see section 11)	350.00	10000.00
	Mary Thompson	U.S. Geological Survey-Southern Appalachian Information Node - Knoxville, TN	various (see section 11)	15.00	150.00
	Nora Schubert	Southern Appalachian Highlands Conservancy - Johnson City, TN	various (see section 11)	1000.00	10000.00
TOTALS				**3375.00**	**62150.00**
Secondary Contributors					
Metadata Compilation	Nora Murdock	National Park Service - Asheville, NC	correspondence, document compilation	2.00	In-kind services
	Rob Sutter	The Nature Conservancy - Durham, NC	correspondence	0.15	In-kind services
	Laura Mansberg-Cotterman	University of North Carolina Herbarium, Chapel Hill, NC	correspondence	0.15	In-kind services
	Alan Weakley	University of North Carolina Herbarium, Chapel Hill, NC	correspondence	0.15	In-kind services
	Marj Boyer	North Carolina Plant Conservation Program, Rougemont, NC (retired)	correspondence	0.15	In-kind services
	Gary Kauffman	U.S. Forest Service - Highlands, NC	correspondence	0.15	In-kind services
	Jame Amoroso	North Carolina Natural Heritage Program, Raleigh, NC	correspondence	0.50	In-kind services
	Dave Danley	U.S. Forest Service - Burnsville, NC	correspondence	0.50	In-kind services
	Karin Heiman	private contractor - Asheville, NC	correspondence	2.00	In-kind services
	Alan Smith	Mars Hill College - Mars Hill, NC	correspondence, document compilation, meetings, field work	20.00	In-kind services
	Paul Somers	Massachusetts Division of Fishies and Wildlife - Westborough, MA	correspondence, document compilation	7.00	In-kind services
	Paul Hamel	Southern Research Station - Stoneville, MS	correspondence	3.00	In-kind services
	Michael Schafale	North Carolina Natural Heritage Program, Raleigh, NC	correspondence, meeting, document compilation	3.00	In-kind services
	Chris Ulrey	National Park Service - Asheville, NC	correspondence	3.00	In-kind services
	Steve Simon	U.S. Forest Service - Asheville, NC	correspondence, meeting	4.00	In-kind services

	David McFee	U.S. Forest Service - Burnsville, NC	correspondence	3.00	In-kind services
	Darlene Kucken	North Carolina Divisin of Water Quality - Raleigh, NC	correspondence	2.00	In-kind services
	Hank Gamble	volunteer - Southern Appalachian Highlands Conservancy; Johnson City, TN	correspondence	0.15	In-kind services
	Paul Bradley	U.S. Forest Service - Burnsville, NC	correspondence	0.50	In-kind services
	Misty Franklin	North Carolina Natural Heritage Program, Raleigh, NC	correspondence	0.15	In-kind services
	Julie Judkins	Appalachian Trail Conservancy - Asheville, NC	correspondence	1.00	In-kind services
USFS plan					
	Joe McGuiness	U.S. Forest Service - Unicoi, TN	correspondence, meeting, document compilation	0.50	In-kind services
	Steve Simon	U.S. Forest Service - Asheville, NC	correspondence	0.50	In-kind services
	Terry Seyden	U.S. Forest Service - Burnsville, NC	correspondence	0.50	In-kind services
	Michael Schafale	North Carolina Natural Heritage Program, Raleigh, NC	correspondence, meeting	0.15	In-kind services
Archiving					
	Jamey Donaldson	private contractor - Shady Valley, TN	correspondence, document compilation, meeting		In-kind services
	Norma Myers	Eastern Tennessee State University - Johnson City, TN	correspondence	6.00	In-kind services
	Tom Blevins	Mount Rogers National Recreation Area, Mount Rogers, VI	correspondence	1.00	In-kind services
	Kristen Johnson	National Park Service - Gatlinburg, TN	correspondence	1.00	In-kind services
	Nature Bib personnel	National Park Service - CO	correspondence	1.00	In-kind services
	Beth Bockoven	The Nature Conservancy - Saluda, NC	correspondence, document identification	1.00	In-kind services
	Tom Laughlin	Eastern Tennessee State University - Johnson City, TN	correspondence	0.10	In-kind services
	Tim McDowell	Eastern Tennessee State University - Johnson City, TN	correspondence	3.00	In-kind services
	Mary Fanslow	volunteer - Southern Appalachian Highlands Conservancy; Kingsport, TN	correspondence	10.00	In-kind services
	Joe McGuiness	U.S. Forest Service - Unicoi, TN	correspondence, document compilation	0.50	In-kind services
GIS					
	Laura Pickens	U.S. Fish and Wildlife Service - Asheville, NC	correspondence, identification of GIS resources	50.00	In-kind services
	UT GIS personnel	University of Tennessee - Knoxville, TN	correspondence, digitizing	3.00	In-kind services
	ETSU GIS personnel	Eastern Tennessee State University - Johnson City, TN	correspondence, meeting	2.00	In-kind services

	Terry Giles	U.S. Geological Survey - Fort Collins, CO	correspondence	3.00	In-kind services
Administrative Overhead	Megan Sutton	Southern Appalachian Highlands Conservancy - Asheville, NC	correspondence, software compilation	3.00	In-kind services
Metadata Records	Kristy Urquhart	Southern Appalachian Highlands Conservancy - Asheville, NC	correspondence, general administrative overhead	30.00	In-kind services
	Cheryl Solomon	U.S. Geological Survey - Lanham, MD	correspondence, compilation of final metadata records	21.00	In-kind services
TOTAL				**189.80**	**0.00**
Associated Costs					
	Aerial photography processing and software	NA	NA	NA	2100.00
	Toolset expansion equipment	NA	NA	NA	2442.00
	Supplies	NA	NA	NA	282.00
	Travel	NA	NA	NA	918.00
	Indirect administrative support	NA	NA	NA	500.00
TOTAL				**0.00**	**6242.00**
GRAND TOTALS				**3564.80**	**68392.00**

Appendix C – Summary of Satellite Imagery Acquisition Research for Monitoring and Management of the Roan Grassy Balds

The use of satellite imagery for assessing change in the spatial extent of the balds as well as for analyzing the effects of various management practices is a possible option for future management. During the FY05 project period, the various options available for satellite imagery were examined so as to determine what might work best. Though aerial/airplane imagery would work for this type of application, the likelihood of performing time-change analysis with imagery of similar specifications is much greater with satellite imagery. Since the management plan for the balds has yet to be defined in a way that incorporates integrated science and specificity, the exploration of imagery was pursued in a manner that identifies 1.) some of the questions that should be addressed before imagery is acquired and 2.) some of the various types available.

The justification for acquiring imagery should come out of a full assessment of management needs, goals, and objectives. Digital imagery is not cheap, so the latter aspects should be fully assessed so as to maximize the use of funding resources. In this case, the assumed purpose of monitoring change in the spatial extent of the balds due to the encroachment of woody plants was pursued, digital images with mid-infrared wavelengths would be necessary to differentiate the various types of vegetation. At the very least, three bands of data are needed: visible green, visible red, and near infrared. This type of imagery is also known as color-IR imagery. Also, the state of the imagery (raw, georeferenced, orthorectified, etc.) is an important consideration. Due to the extreme relief of the Roan Mountain area, orthorectified imagery should be used so that excessive error is not introduced into the analysis. This is particularly important for analyzing change in spatial extent over time.

The scale and resolution appropriate for this analysis is also a critical driving factor behind which types of imagery should be used. The primary question with this is: At what level does monitoring need to occur (30 meter resolution, 10 meter resolution, 1 meter or better resolution etc.) to determine change over time and the effects of on-the-ground adaptive vegetation management? Ideally, imagery with resolution better than 1 meter would work the best for assessing the change over time. However, with resolution comes the important real consideration of cost since resolution directly correlates with cost. It is important that the availability of funding for future acquisition and future analysis should be fully taken into consideration so that similar imagery can be acquired at a later date for temporal analysis. The availability of expertise for working with such imagery, as well as the availability of imagery analysis software such as ERDAS IMAGINE should also be considered so that the imagery can be utilized to the fullest extent possible. Pinpointing the time of year that would allow for the greatest visual variation between plant types on the balds, as well as taking into consideration the timeframe of the various current management practices such as mowing and cutting, is also critical for determining when data should be acquired.

It was suggested that Landsat imagery would work for the purpose of analyzing general change in spatial extent on the balds, but further exploration of this would be needed to verify this recommendation. The resolution of Landsat is likely not high enough for identifying specific plant species, but could work well for looking at change in woody/non-woody plant boundaries. The majority of Landsat data has 30 meter resolution, except for the Landsat ETM+ imagery which has a 15 meter panchromatic band. Testing Landsat's applicability to this would be feasible since some Landsat imagery can be obtained, already orthorectified, for free from http://earthexporer.usgs.gov. After performing queries for imagery that covers all of the balds, a couple datasets were discovered that could be used for testing this option. Free orthorectified Landsat ETM+ data is available for 2000/05/16 and 2001/10/26, as well as Landsat TM orthorectified imagery available for 1988/09/21 and 1987/06/06. Landsat MSS orthorectified imagery is also available for free from 1976/06/06 and 1973/10/27. Landsat 5 TM would compare with ETM+ when looking at change over time. However, TM data does not have the panchromatic band that ETM data does, so if that band is to be used with ETM+ data, it isn't available with TM data. Landsat 1-5 MSS data can also be used for time change analysis with other Landsat types, but the resolution for MSS is lower than TM and ETM+.

Many other dates for these three are available for purchasing, however the status of the imagery (orthorectified or not, etc.) might vary and in turn influence cost. If in fact Landsat imagery is determined usable by testing the free imagery mentioned above via classifications with imagery processing software, Landsat data could prove to be an excellent means for general time change analysis since Landsat data is available back through the 1970's. Landsat is also much cheaper than some of the other widely-used commercial types with regards to the feasibility of acquiring similar imagery in the future.

More details about Landsat, as well as other available satellite data products at
http://edc.usgs.gov/products/satellite.html
http://landsat.usgs.gov/project_facts/history/

Landsat imagery and other data can also be acquired from the University of Maryland Global Land Cover Facility:
http://glcf.umiacs.umd.edu/index.shtml

Other options for satellite imagery acquisition include a number of private companies that provide very high resolution imagery in many different formats, resolutions, etc. Such imagery with resolutions of 1 meter or better would be excellent for differentiating specific types of vegetation on or around the balds. However, many of these companies have a minimum order size and typically are much more expensive than other options such as Landsat. As mentioned, this is an important consideration with regards to future imagery acquisition and available funds. These companies include:

SPOT imagery: http://www.spot.com/html/SICORP/_401_.php
IKONOS imagery from Space Imaging: http://www.spaceimaging.com/products/ikonos/index.htm
Quickbird imagery from Digital Globe: http://www.digitalglobe.com/about/quickbird.html